LEVEL II

(12)

THEATER-LEVEL GAMING AND ANALYSIS WORKSHOP FOR FORCE PLANNING

Volume I — PROCEEDINGS

September 22-29, 1977, held at Xerox International Center for Training and Management Development, Leesburg, Virginia

SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

LEVEL II

(6)

THEATER-LEVEL GAMING AND ANALYSIS WORKSHOP FOR FORCE PLANNING

Volume I — PROCEEDINGS

September 22-29, 1977, held at Xerox International Center for Training and Management Development, Leesburg, Virginia

(12) 322p.

(11) 1977

Sponsored by
OFFICE OF NAVAL RESEARCH
Code 431
Arlington, Virginia 22217

Chairman:
(10)
Lawrence J. Low
SRI International
Menlo Park, California

(15)
Contract No. N00014-77-C-0129

411 281

Contents

Partial Contents: Gaming Utility from the User's Viewpoint; Status of Theater-level Simulation Models, Present and Future, and Problems of Model Structure;

Contents (Cont'd)

Illustrations

Illustrations (Cont'd)

Illustrations (Cont'd)

Illustrations (Cont'd)

Illustrations (Cont'd)

Foreword

The "Theater-Level Gaming and Analysis Workshop for Military Force Planning" meeting was held at the Xerox International Center for Training and Management Development, Leesburg, Virginia, on 27-29 September 1977. Volume I documents the proceedings of this meeting. Volume II presents an "author's-eye view" of the accomplishments and the problem issues associated with this field of endeavor, and further outlines a concept for research that addresses these problem issues.

Volume I has been prepared from edited transcripts of tape recordings made at the meeting. As a consequence, the informal, colloquial style that characterized the presentations, panel discussions and exchanges among participants has been adhered to. When a presentation was based on highly technical material, the paper from which it was derived is presented as an appendix to Volume I. In all but six cases, the formal presentations made at the meetings have not been subjected to a review by the speakers who made them. The decision to short-circuit the review process stemmed from the large number of presentations made and a strong desire on the chairman's part to publish the report of the meeting before the onset of senility.* As a consequence, I assume full responsibility for any misinterpretations, transgressions, errors, or inaccuracies that have resulted and fervently hope that they are minimal.

Assistance to the author in structuring the meeting and selecting participants was provided by a Steering Committee composed of the following individuals:

Dr. Jerome Bracken, Institute for Defense Analysis

Prof. Toke Jayachandran, on assignment to Office of Naval Research from Naval Postgraduate School

Prof. William Lucas, Cornell University

Prof. Michael Sovereign, Naval Postgraduate School

Dr. Thomas Varley, Office of Naval Research.

To these people I would like to express my deep appreciation for their efforts and most particularly to Drs. Bracken and Sovereign for the assistance they have provided since the meeting.

The list of those who graciously gave of their time would be incomplete without mention of Prof. James Taylor of the Naval Postgraduate School; Herbert K. Weiss of Litton Industries; Dr. Frank Kapper, SAGA; and Carl Eulenstein, CINCPAC Staff; and I also wish to thank LCOL Robert Doty USAF, CINCPAC Staff for preparing the theater-level gaming geneology chart in Volume II of this report. Finally, there is a debt of gratitude owed to my colleague Phyllis Dorset, which I can hardly ever hope to repay, for the staggering task of editing all of the material for this report. Appreciation is also due Luba Wittke, Ruth Avery and Varena Strandness and their associates for their efforts in illustration, composition and layout.

*There are those who might argue that it is already too late.

Introduction

Nature of the Problem

The past several years have witnessed a definite kindling of government interest in theater-level gaming of general purpose forces. A relatively small but select number of practitioners in gaming have responded to diverse specific demands from the defense establishment and other agencies for models and analyses in a way that can be characterized as single-minded and purposeful with little time, opportunity, or inclination for crosstalk among the community of gamers and users. In addition, this type of gaming has met with mixed user reaction that, at best, can be described as skeptical and cautious in its acceptance. Many users and researchers in the analysis community are critical of attempts made to game at the theater-level with models or simulations because of the known existence of a variety of shortcomings. First and foremost, there is no unifying "theory of combat" to guide model development; there is much concern about the necessary resolution you must have in a model to properly capture the vagaries of battle; there is uncertainty about the validity of attrition algorithms that are used and both the applicability and quality of data that constitute inputs to the models; finally, there is general dissatisfaction with model complexity and lack of transparency and with the fact that they are costly to develop and operate. Validation of the models and the data they require is believed by many to constitute a rather formidable research and/or operational undertaking; yet, this appears to be one of the most pressing problems.

Added to all of this is the ever-increasing awareness of a need to bring man back into the gaming loop (after his removal in earlier simulation efforts), either through man-computer interactive techniques or through mathematical optimization techniques such as game theory. This comes about because of the obvious two-sided nature of combat with the strategic and tactical decisions that must be made on either side in the light of perceived activities of the opponent. The application of game theory-techniques to this class of problem is somewhat imperfect, it is complex and subtle as well as difficult to implement. There is considerable disagreement among researchers in the area about the extent to which one can make simplifying assumptions and approximations and still arrive at anything approaching valid answers.

Granting the resolution of these issues, there still remains the difficult task of communicating an understanding of these techniques, and the results obtained from them, to the man who must make the decisions. Such understanding is necessary if the quantitative input to the decision process is to be placed in proper perspective with the many qualitative inputs. In sort, the field of endeavor is in some disarray, characterized by fragmented, compartmentalized efforts with relatively little information exchange among investigators or user activities.

On the plus side, however, is the fact that these gaming techniques afford a powerful potential for dealing with entire classes of problems that heretofore were unapproachable in a consistent, structured manner: problems of net assessment and the evaluation of arms control and disarmament alternatives: problems of force structure and force level planning. Encouragingly enough, there are those in key decision making positions who believe that any rational method or construct that enables one to come to grips with these problems is a useful adjunct to, or perhaps better than, older methods of arriving at force postures and balances purely through the use of static indicators or "military judgment." And, needless to say, there are some who do not share this viewpoint.

It is for these reasons that the Naval Analysis Program of the Mathematical and Information Sciences Division of the Office of Naval Research sponsored the subject conference and workshop. The purposes of the meeting were to foster better communication between expert members of the user community and the applied and basic research communities, to explore ways of solving the significant array of problems associated with theater-level gaming, and to formulate eventually a logical, orderly plan for continuing research in the area.

Observations on Gaming

The following material was prepared to define some terms, explain some concepts, and to establish, generally, the train of logic from which the structure of the meeting evolved.*

Gaming is jointly nested in military science and operations research, areas that are inordinately broad in scope and that contain elements that have been the object of intensive and expanding interest in recent years. The rapidity of this expansion of interest among various activities of the three services, the JCS, the DoD, the intelligence community and the arms control interests in government, has led to inevitable difficulties with terminology. The ambiguities and lack of consistence and precision in terminology, in turn, have led to varied interpretations of the relationships between gaming and other forms of operations analysis and between the different types of games themselves. Some excellent work in refining terminology related to gaming has already been performed.† The discussion that follows conforms to these previous efforts and attempts to accommodate some very recent concepts in gaming.

Most of the problems associated with the definition of terms in gaming are the result of the careless drift of the words "games" and "gaming" into the military operations research vocabulary. Webster defines a game as "a situation involving opposing interests given specific information and allowed a choice of moves with the object of maximizing their wins and minimizing their losses." While this definition may have a game theory ring to it, it should be noted that it most certainly applies, in general, to warfare. Noting further that "analytic models, machine simulations and games" are often strung together in prevailing attempts to classify the analytic techniques that are used to solve two-sided military problems, it can be argued that this form of classification mixes apples and oranges. Models and simulations are indeed *techniques*, while games are related to *situations* and *behavior*. If a game is considered to be any type of analysis or modeling effort that contains an explicit representation of two (or more) sides in defining an adversary or conflict situation‡, then a significant step would have been taken toward clarifying gaming terminology. This definition of a game should apply whether neither side, one side, or both (or more) sides are afforded a choice of moves in the mathematical formulation of the problem.

In clarifying the terminology, it is useful to review briefly the fundamentals of combat. Figure 1 shows the basic concept of combat. Simply stated, all combat involves the interaction between opposing forces, designated RED and BLUE. These forces are composed of *men* and *equipment*, are governed by *operating procedures*, and involve some measure of *combat support*. Both forces function in an *operational environment*, which is composed of natural factors such as weather and terrain. The interaction between RED and BLUE results in a *combat operations outcome*, which can be measured in a variety of ways.**

A morphological matrix can be constructed in three dimensions to review the field of military gaming in all of its forms. The dimensions are concerned with:

1. Gaming and analysis *technique*
2. Gaming and analysis *scope*
3. Gaming and analysis *application*.

---

*This material is reproduced from a working paper entitled "Theater-Level Gaming and Analysis Workshop for Military Force Planning (Concept and Plan)" May, 1977. The paper was submitted to all invitees a month or two prior to the meeting in order to define its scope and content.

†M. Shubik and G. D. Brewer, "Models, Simulations and Games — A Survey," Rand Report R-1060-ARPA/RC (May 1972), UNCLASSIFIED.

S. Bonder, "An Overview of Land Battle Modeling in the U.S.," Proceedings of the 13th U.S. Army Operations Research Symposium (November 1974), UNCLASSIFIED.

‡The writer recognizes that many analyses are concerned with "suboptimization" problems involving man-machine system performance in military operations environments where the existence of an adversary is only implied. Whenever the existence of an adversary does not enter explicitly into the definition of the problem, the solution techniques would not, for purposes of this dicussion, qualify as a game.

**Three basic classifications of combat outcome suggest themselves

Annihilation: the forces of one side are destroyed virtually en toto on the battlefield by those of the other side. Vanquished force remnants are routed, captured, or surrender to the enemy.

Territorial conquest: The seizure (capture and occupation) of all of one side's territorial objectives; hostilities are terminated by the route, capture, or surrender of opposing forces.

Stalemate: the achievement of objectives and/or the number of casualties suffered lead to a protracted conflict or a negotiated settlement.

The first two, defining win/lose situations, can be used singly or in combination; the third defines a draw.

These dimensional categories can be further expanded as follows:

1. Gaming and analysis technique
   a. Military exercises (field, fleet, air, joint)
   b. Manual war games
   c. Computer assisted manual war games
   d. Analytic/computer games (analytic models, simulation, optimization)
   e. Interactive computer games
2. Gaming and analysis scope
   a. Theater-level conflict
   b. Major general engagement or battle (in-theater)
   c. Local engagement, "many-on-many units"
   d. Local engagement, "one-on-one/many units(s)"
3. Gaming and analysis application
   a. Force planning
   b. R&D planning, management, and evaluation
   c. Operational planning and evaluation
   d. Training and education.

Some of the descriptors for these techniques are sufficiently broad (and perhaps obscure) to require brief amplification. Others, at least for this overview, are largely self-explanatory. Perhaps the most critical dimension from the standpoint of reconciling the different usages of familiar terms in the operations research and analysis community is that of gaming technique. A few comments about each of the techniques follow:

Gaming and Analysis Technique

(1a) Military Exercises (Field, Fleet, Air, or Joint) — These are exercises employing men and equipment that might actually be used in battle in the near term. Realism in nearly all aspects of combat shown in Figure 1 is the

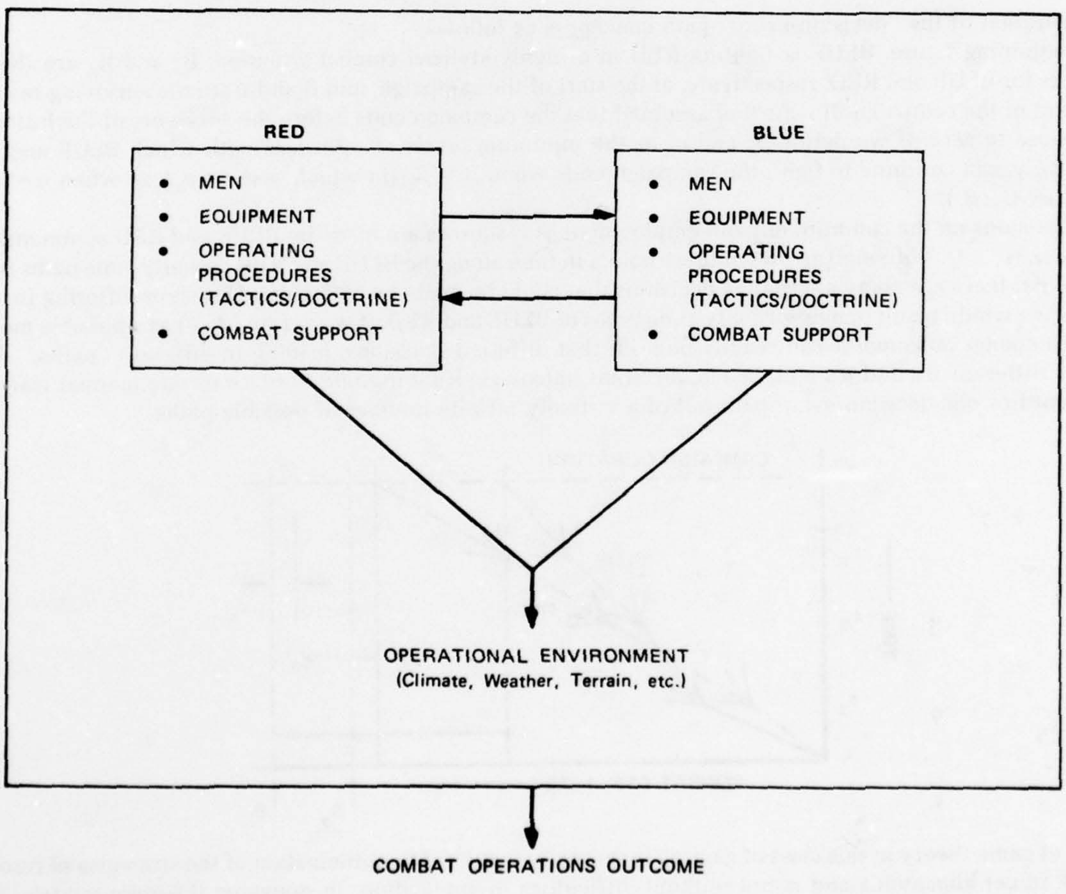


FIGURE 1 A CONCEPT OF COMBAT

strongest feature of this form of gaming. These exercises are used largely for training, operational evaluation of current equipment and man/machine interfaces, for development and evaluation of tactics, and as a measure of combat readiness. Relatively speaking, they are not only very costly but are limited in that they cannot, by their very nature, accommodate future operational systems. Then, too, the commanders, personnel, and equipment for the RED (enemy) forces in these exercises are actually drawn from BLUE (friendly) resources. This can be prejudicial to the way in which RED operates, introducing potentially significant biases into the experiment.

(1b) Manual War Games — These are also referred to as free-form games and involve BLUE and RED teams and a referee group operating within the framework of a scenario and generally played around a map. They are perhaps the least costly form of gaming and afford the opportunity to deal with future systems and force concepts, provided that these can be fairly well defined. Perhaps the most significant shortcoming of manual war games (and the military exercises discussed above) for particular applications is that the outcome of a play of the game is generally unique and largely irreproducible, depending heavily on which of a multiplicity of "decision/event paths" is taken by the BLUE and RED commanders and the tides of battle during the game play.* Nonetheless, manual war games have enjoyed widespread use in the training of military combat commanders. Furthermore, they provide insights into the structure and salient features of a particular combat situation and a feel for the *kinds* of systems and forces that are needed to achieve specified objectives. However, they are particularly awkward to use and inappropriate for addressing questions of "how many."

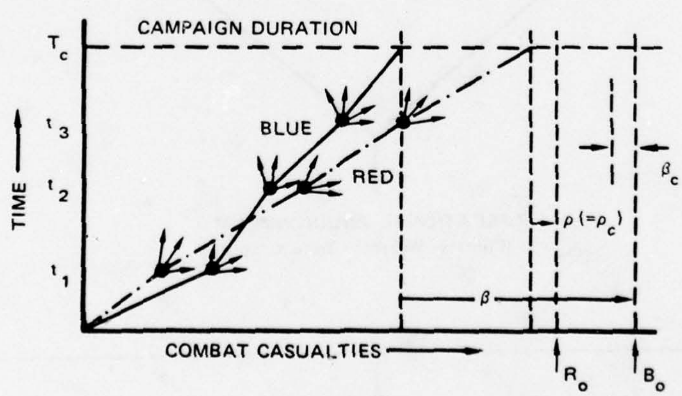
(1c) Computer-Assisted Manual War Games — These are basically the same as (1b) above except that computers and computer models/software are used to assess the outcome of command decisions, which places a lesser requirement on the need for referees. Many of the more routine types of engagements that occur in a military operation or campaign have been automated (computerized) to standardize to a degree the assessment of these military encounters. By and large, this form of gaming, while more costly than (1b) because of the involvement of computers and computer software, has the same basic advantages and disadvantages as those attributed to manual war games. These games represent a class of what are sometimes referred to as man-machine games in the literature (see also 1e).

---

*An illustration of the "decision-event" path concept is as follows:

In the adjoining figure, BLUE is fighting RED in a highly-stylized combat situation. B_0 and R_0 are the total resources for BLUE and RED, respectively, at the start of the campaign, and β and ρ are the surviving resources at the end of the campaign. It is further assumed that the campaign ends before the survivors of the losing side are reduced to zero. If we define β_c and ρ_c as the minimum levels of resources with which BLUE and RED, respectively, can continue to fight, the campaign ends when $\beta = \beta_c$ (in which case $\rho \geq \rho_c$), or when $\rho = \rho_c$ (in which case $\beta \geq \beta_c$).

Major decisions on the commitment and employment of resources are made by BLUE and RED commanders at times $t_1, t_2, t_3 \ldots t_c$. The small arrows at these points in time along the BLUE and RED casualty-time paths reflect the fact that there are many alternative decisions that could be made by both sides. Decisions differing from the ones shown would result in new casualty-time paths of BLUE and RED. If we define $(\beta - \rho)$ as a possible measure of the campaign outcome, it can readily be seen that different decisions, leading to different "paths," would result in different β's and ρ's and, hence, different outcomes. Running one exercise or one manual war game only identifies *one* decision/event path out of a virtually infinite number of possible paths.



†The use of game theory in this class of gaming is generally restricted to optimization of the strategies of resource (military force) allocations and is not without difficulties in application. In nongame theoretic models, these strategies are fixed for both BLUE and RED.

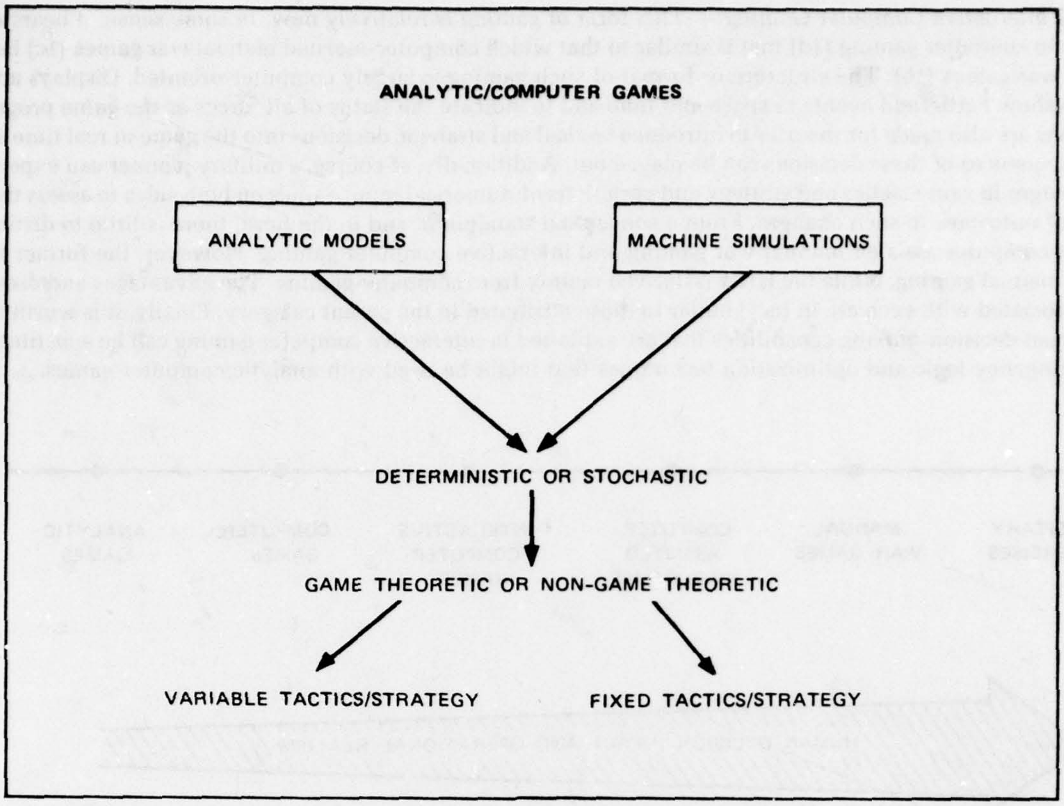


FIGURE 2 TYPE CLASSIFICATION OF ANALYTIC/COMPUTER GAMES

(1d) Analytic/Computer Games — In this category of gaming, the human being is removed from the actual play or execution of the game. Instead, analytic models and machine simulations are the analysis devices or techniques used exclusively to construct an abstraction of the combat situation. These models and simulations may be deterministic or stochastic. Furthermore, they may or may not be game-theoretic[†] in character (see Figure 2). This class of gaming delves more deeply into abstractions, particularly in analytic gaming. What is sought in return through an attempt to express all abstractions in mathematical terms or computer language is a more definitive "outcome" expectancy, as well as a better understanding of the complex structure of the combat process. This class of gaming would appear to be particularly useful for applications such as force planning. In reference to an earlier argument, Figure 2 shows analytical models and machine simulations used as analysis devices or techniques to solve analytic/computer games. Models and simulations are more appropriately subordinate to rather than coequal with games in the gaming definition hierarchy.

While analytic models are highly abstract (which sharply limits their application), they are nonetheless very effective in providing important insights into the relationship among basic combat parameters. These insights can be used to advantage in computer gaming or simulations. Game simulations, on the other hand, attempt to "act out" the combat process, and the computational speed and memory of the digital computer permits the development of more complex detailed and realistic models with this technique than can be handled by closed form, analytic techniques. Simulations are generally more costly than manual war games and certainly more so than analytic game models. Furthermore, the apparent convenience and speed in problem solving associated with simulation techniques are offset to varying degrees by the time involved in preparing the input information required for the gaming model. Of course, once this information has been assembled, many cases of particular interest can be run and analyzed with relative speed and ease.

Conceptually, this class of gaming recognizes that man and his decisions remain a very important element of warfare. An attempt is made to accommodate this concept mathematically through the use of contingency logic (in computer games) that governs the way a commander will employ his forces, depending on what occurs during the battle, or through the use of optimization techniques such as game theory to define the optimal tactics and strategies to be employed by both sides. Much of the attempt to mathematically accommodate man remains as an area for further research.

(1e) Interactive Computer Gaming — This form of gaming is relatively new. In some sense, it bears a relationship to computer gaming (1d) that is similar to that which computer-assisted manual war games (1c) bears to manual war games (1b). The structure or format of such gaming is largely computer-oriented. Displays are provided to show battlefield events in space and time and to indicate the status of all forces as the game progresses. Provisions are also made for the user to introduce tactical and strategic decisions into the game in real time so that the consequences of these decisions can be played out. Additionally, of course, a military planner can experiment with changes in game tactics and strategy and variations of numerical input values on both sides to assess the sensitivity of outcomes to such changes. From a conceptual standpoint, and in the limit, there is little to distinguish between computer-assisted manual war gaming and interactive computer gaming. However, the former has its roots in manual gaming, while the latter is derived mainly from computer gaming. The advantages and disadvantages associated with each are in fact similar to those attributed to the parent category. Finally, it is worth noting that human decision-making capabilities that are exploited in interactive computer gaming can be substituted for the contingency logic and optimization techniques that might be used with analytic/computer games.

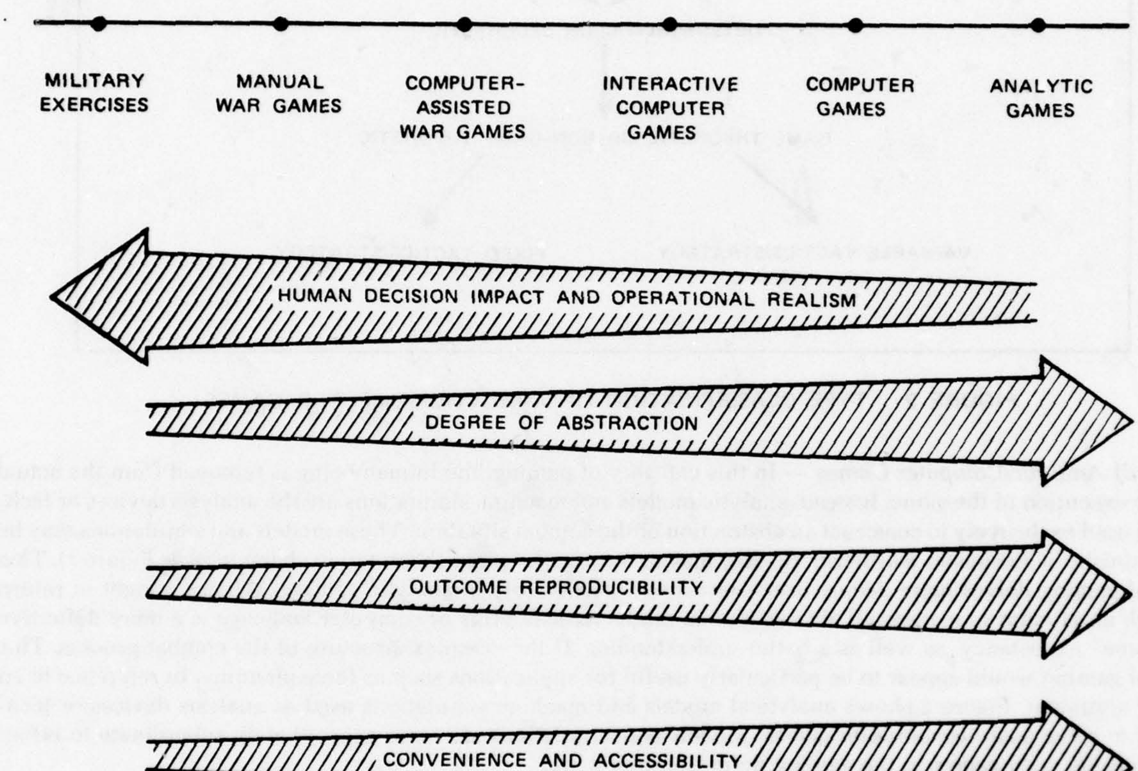


FIGURE 3 GAMING SPECTRUM AND GAME CHARACTERISTIC TRENDS

By way of summary, the spectrum of gaming is presented in Figure 3 along with a schematic of the relationships among key characteristics that are generally associated with gaming. The costs associated with various gaming techniques, however, do not follow such an orderly ascending/descending pattern. A relative cost-ranking in descending order for the techniques that have been discussed, fixing the problem scope, would generally be as follows:

1. Military exercises
2. Interactive computer games
3. Computer games
4. Computer-assisted manual war games
5. Manual war games
6. Analytic games

Gaming and Analysis Scope

There is perhaps less need for terminology amplification in this dimensional category than in the previous one. However, a few comments are offered to more clearly relate the descriptors used in the breakdown of this category to the types of conflict situations that are commonly modeled.

(2a) Theater-Level Conflict — This level of conflict treats warfare that involves an entire geographic area of the globe in which ground, sea, and air forces may become directly employed in operations and includes the theater of operations and the zone of the interior. In this level, not only are the warring forces themselves considered but so are all of their support in the form of ground, sea, and air lift for initial deployment, reinforcement, supply and resupply. Strategic matters such as mobilization build-ups and the forward basing of forces also must be considered, as appropriate. Global warfare, a yet higher level of conflict, can generally be treated in terms of the involvement of two or more theaters of operations. Gaming at these levels is highly aggregated.

(2b) Major General Engagement or Battle — As the name implies, a major engagement, in-theater, involves ground, sea, or air forces singly or in combination. In gaming these engagements, combat operations are emphasized, and combat support is generally treated implicitly or exogeneously to the exercise.

(2c) Local Engagements, "Many-on-Many Units" — This category is a catch-all for local engagements that involve multiple organizational or systems units on both sides. The factors spelled out in Figure 1 still apply here and, at this level of combat, are considered in greater detail than in (2a) or (2b) with the exception of combat support, which is largely implied or assumed. Units belonging to either side may be homogeneous or nonhomogeneous as to type or class, and opposing units may be of similar or dissimilar types and classes. Traditionally included in this category are, for example, situations involving infantry units against tanks, a naval task force defending itself against attack by enemy submarine-launched cruise missiles, or tactical air units operating in close support of ground forces.

(2d) Local Engagement, "One-on-One/Many Unit(s)" — This low level in the hierarchy examines the combat performance or effectiveness of a single organizational or system unit against some reasonable level or portion of the threat. The threat, in turn, may consist of a single unit of a certain type or class or of multiple units of the same or differing types and classes.

Another candidate element warrants mention because it is used all too frequently, yet it does not qualify as a game under the conditions specified here. This element is the use of a direct comparison of the performance parameters and characteristics of two opposing systems (often, hardware systems, but not necessarily restricted to these) with some rather subjective weighting factors applied to certain parameters that reflect a decision maker's value judgment of what is really important in the exercise. There is not formal analysis structure involved, and it is the utter simplicity of the technique that makes it persistently attractive. Its validity is highly questionable for any use beyond establishing some very crude measures of relative performance. Deeper, but nonetheless fundamental, questions of the type "Is faster better?" must be answered by more complex methods of analysis and gaming.

In the progression from (2d) to (2a), it is obvious that higher and higher levels of aggregation of structure and information are forced into the gaming procedure. Conversely, the characteristics of detail and resolution increase in going from (2a) to (2d).

Gaming and Analysis Application

Gaming and analysis application is largely self-explanatory with only a word or two of amplification needed.

(3a) Force Planning — Force planning includes planning directed toward establishing force structures (organizations and equipment), force mixes (among force organizational entities), and force levels (the sizing of the various force entities) to maximize the U.S. military posture under prevailing budget constraints.*

(3b) R&D Planning, Management, and Evaluation — In this area of application, gaming is chiefly concerned with the analysis of hardware systems (weapons, sensors, vehicular platforms, and the myriad supporting

---

*To clarify the usage of terms in a somewhat troublesome area, the following definitions are advanced:

Force structure/composition — The organizational form and the types of equipment associated with land, sea, and air combat units of specified size in the traditional military sense (e.g., battalion, carrier task group, tactical air wing, etc.) and with their support elements.

Force level — The number of combat units of the same type (of specified composition) that constitute a significant element of a larger force assigned to a military purpose.

Force mix — The numerical mix of, or ratio between, land, sea, and air units that, in combination, constitute a force assigned to a military purpose.

Force planning, as already indicated, is involved with all of the above.

systems to be found in the defense establishment) and covers aspects of systems conceptualization, systems design and development, systems evaluation and systems selection. This element includes Operational Test and Evaluation (OT&E). The element encompasses the major bulk of gaming work performed today.

(3c) Operational Planning and Evaluation — This is gaming used in the design of military exercises and the development and evaluation of operational concepts, tactics, and doctrine.

(3d) Training and Education — This element is concerned with the training of military commanders at various levels in the art of decision making under conditions that are expected to prevail in combat or wartime as well as the education of students of military science in the fundamentals of warfare, while introducing them to the principles of analysis and gaming.

In the breakdown of the gaming and analysis application dimension, there is implied a hierarchy, of sorts, in the importance of the decision making that is supported by gaming. That this hierarchy is an ordering of activities that goes from "most important" at the top (3a) to "least important" at the bottom (3b) might well be argued against since all activities shown are necessary to the operation of a military establishment. However, in practice, the decision makers concerned with the activities listed tend to conform to a "top-to-bottom" hierarchical pattern when one takes into consideration their positions in the DoD.

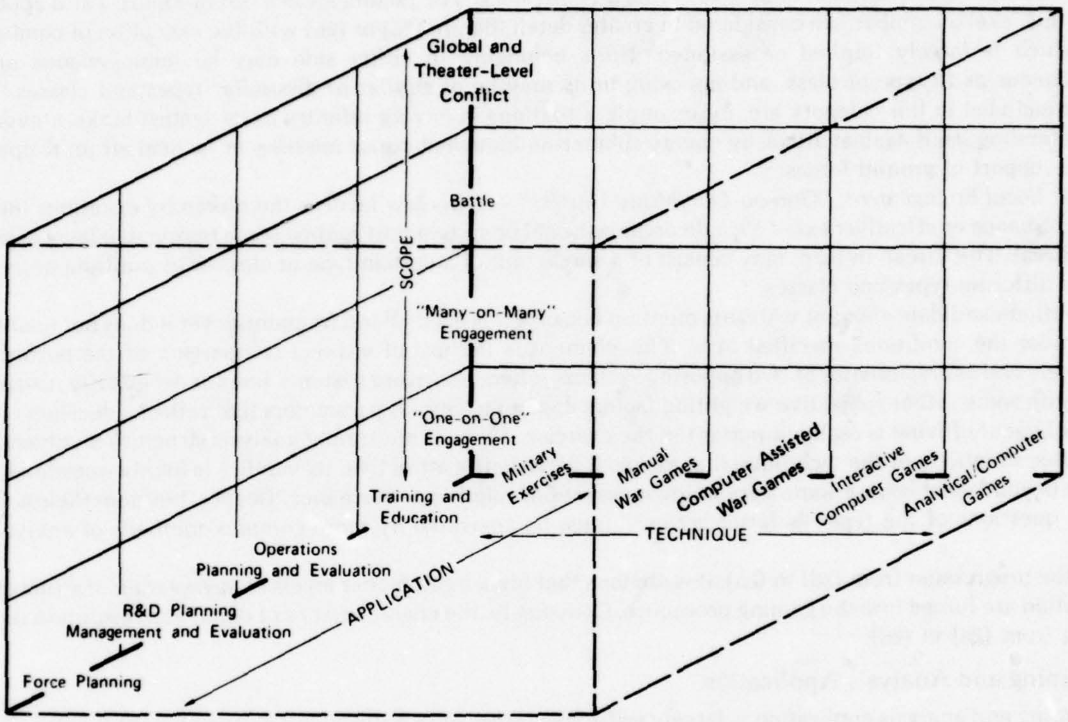


FIGURE 4 GAMING CLASSIFICATION MATRIX

Figure 4 shows the matrix of gaming. The amount of analysis activity in the United States today that it represents is enormous indeed. Clearly, any research to be performed in gaming must be selective and focused on particular elements of this matrix if the effort is not to become untenable in its proportions.

The selection of military force planning as the gaming application of interest for this conference is not difficult to justify. It is a matter of concern at the highest levels of the defense establishment, it affords opportunities for some exciting, challenging research at the fundamental and applied levels, and it presents some interesting philosophical and technical problems in generally enhancing the credibility of quantitative methods at high levels of decision making. As areas of endeavor go in the field of OR/SA, the attention given the force planning problem has been relatively minimal so that research performed in this area should have significant impact.

With the selection of force planning as the appropriate application topic on which to focus attention, the other two dimensions of the matrix are rapidly locked into analytic/computer games for *technique* and theater-level conflict for scope. Considerable extension into interactive computer games and computer-assisted games is, of course, possible. In any event, primary concern in this conference and workshop is biased toward the upper, forward, righthand corner of the matrix cube shown in Figure 4.

Meeting Structure

With force planning as the underlying applications objective for the gaming conference, quantitative techniques are necessary if the effects of varying force structures or composition and force levels are to be *measured* in some definitive way. With such techniques, we also seek high levels of outcome reproducibility (Figure 3) while hopefully doing as little mischief as possible to conflict with realism in our abstractions. Although briefly stated, it is this reasoning that leads us to the previously discussed analytic/computer games as the most promising gaming technique for this type of application. Furthermore, because of the inherent limitations of analytic games, it is finally computer gaming that should command most of our attention.

Several distinct steps are involved in applying computer gaming techniques to a problem of any sort. Beyond the important step of defining the problem at hand, we have model selection, modification, or development; scenario construction; input data collection and preparation; computer selection and utilization; and analysis of results. This procedure may lead to a variety of iterations involving parts of the process, at least, to explore the sensitivity of outcomes to input variations. The critical elements of this process are as follows:

- The user/decision maker and the problem to be solved.
- The gaming model, its flexibility of structure, and its adaptability to the user's problem.
- The combat modeling techniques employed in the gaming model, their adequacy and validity.
- The data base available for input preparation and its validity.

The topical breakdown of matters that were discussed at the conference/workshop were based on these elements. In addition, game-theory and game theoretical models and their proper role in the process were made part of the agenda. Finally, the important topic of validation, particularly as it applies to the data base and to combat modeling were included in the discussion.

In summary, the following topical designations suggested themselves for the conference sessions:

- User attitudes and concerns
- Gaming model structures
- Combat modeling techniques
- Game theory (mathematical and heuristic optimization)
- Data base and validation problems.

A wide spectrum of research activity is being spanned with the selected subject breakdown shown above. We are dealing with fundamental or basic mathematical research under "Game Theory," applied operations research for the most part under "Gaming Structures and Combat Modeling," possible experimental and historical research under "Data Base" and "Validation," and a measure of psychological or behavioral research when it comes to "User Attitudes."

A series of invited papers were presented in each of the five subject areas identified above. The areas, in turn identified the material content for the five sessions of the meeting. In addition, there were panel discussions in each session following the presentation of papers, with approximately 40% of the total time allocated to the meeting being set aside for such discussions (including a summary session).

Conference/Workshop Objectives

Objectives

The goals of the meeting were to:

- Promote the broadest dissemination among all attendees of the technical discussions that take place and expressed concerns of user activities.
- Resolve problem issues that have been identified to the maximum extent possible and to bring out other related problems or problem areas.
- Produce for each session a panel summary of the status and the research state-of-the-art in the associated technical areas.
- Develop from the discussion and findings of the meeting as a whole, a strategy for responding to the needs for research that have been identified.

Session I: Gaming Utility From the User's Viewpoint

- To present and discuss the types of decisions at various governmental levels that are supported by computer gaming and to identify the more serious deficiencies and shortcomings, as the user sees them, in current practices and techniques.

Session II: Status of Theater-Level Simulation Models, Present and Future, and Problems of Model Structure

- To present salient features of current non-game theoretic simulation models, and of those under development and to identify those existing problems that relate to model structure.

Session III: Theater-Level Combat Modeling Methodology

- To review the state-of-the-art in methodology development for modeling the combat interactions between military forces at the theater-level and for the treatment of related operational phenomena.

Session IV: Game Theory in Theater-Level Modeling: Optimal Solutions and Heuristic Solutions

- To review techniques for the application of game theory and man-computer interactive gaming to theater-level models and to identify some major problems, both theoretical and applied, that require resolution.

Session V: Data Base and Validation Requirements for Theater-Level Models

- To identify efforts, present and future, to obtain and disseminate data required to derive inputs for models of theater-level conflict. To present a review of research directed toward the validation of theater-level combat modeling.

Meeting Agenda

Tuesday, 27 September, Morning Schedule

INTRODUCTION AND WELCOME

Call to Order: Mr. Lawrence Low, SRI International

Sponsor's Remarks: Dr. Thomas Varley, Office of Naval Research

KEYNOTE ADDRESS: "Use of Gaming in Force Planning"

MGEN Jasper Welch USAF, Headquarters, U.S. Air Force, AC/S Studies and Analysis

SESSION I: Gaming Utility from the User's Viewpoint

Chairman: Dr. Frank Kapper, Joint Chiefs of Staff, Studies Analysis and Gaming Agency

Opening Remarks: Dr. Frank Kapper

Theater-Level Gaming in the Formulation of Plans and Policies (Department of Defense)

(1) *Army*: Mr. Philip Louer, Department of the Army, Office of Deputy Chief of Staff for Operations and Plans

(2) *Navy*: Mr. John Shewmaker, Congressional Budget Office

(3) *Planning and Evaluation*: Mr. Robert Schneider, Office of the Secretary of Defense (Planning and Evaluation)

(4) *Net Assessment*: Mr. Andrew Marshall, Office of the Secretary of Defense, Net Assessment

(5) *Joint Chiefs of Staff*: Dr. Frank Kapper

Theater-Level Gaming in the Formulation of Plans and Policies (Defense Related Activities)

(6) *Central Intelligence Agency*: Mr. James Starkey, Central Intelligence Agency

(7) *Arms Control and Disarmament*: COL Miles March USA, Arms Control and Disarmament Agency

DISCUSSION PANEL

Dr. Frank Kapper

Dr. Rex Goad, SHAPE Technical Centre, The Hague

COL John J. Grace USMC, Headquarters, U.S. Marine Crops.

Mr. Frank Hoeber, Consultant

Dr. Hilmar Linnenkamp, Ministry of Defense, Federal Republic of Germany

Tuesday, 27 September, Afternoon Schedule

SESSION II: Status of Theater-Level Simulation Models, Present and Future, and Problems of Model Structure

Chairman: Dr. Jerome Bracken, Institute for Defense Analyses

Opening Remarks: Dr. Jerome Bracken

(8) *Theater-Level Models*: Dr. Seth Bonder, Vector Research Inc.

(9) *A Critique of Four Theater-Level Models*: Prof. Alan Karr, Johns Hopkins University

(10) *Four Model Comparison Study*: LTC Lanny Walker USA, Third Basic Combat Training Center

(11) *Theater-Level Modeling of Conventional, Nuclear and Chemical Warfare*: Dr. Edward Kerlin, Institute for Defense Analyses

(12) *Experiences with Interfacing Ground/Air and Sea War Combat Models*: COL Carl Hess USA, Office of the Assistant Secretary of Defense (Material Acquisition and Logistics)

(13) *Problems of Aggregation and Resolution in Theater-Level Models*: Mr. John Bode, The BDM Corporation

(14) *Outcome, Effectiveness and Decision Criteria for Combat Gaming*: Dr. George Pugh, Decision Science Applications, Inc.

DISCUSSION PANEL

Dr. Jerome Bracken

Dr. David Dare, Defence Operations Analysis Establishment, Great Britain

Dr. Reiner Huber, Hochschule der Bundeswehr München, Federal Republic of Germany

LGEN Glenn Kent USAF (Ret.), Glenn A. Kent Associates

Prof. Michael Sovereign, Naval Postgraduate School

Mr. Ross Thackeray, Ketron, Inc.

Wednesday, 28 September, Morning Schedule

SESSION III: Theater-Level Combat Modeling Methodology

Chairman: Mr. Robert Farrell, Vector Research, Inc.

Opening Remarks: Mr. Robert Farrell

(15) *Attrition of Ground Systems Modeling Overview*: Mr. Robert Farrell

(16) *Attrition of Air Systems Modeling Overview*: Dr. Bruce Anderson, Institute for Defense Analyses
(17) *Attrition of Marine Systems Modeling Overview*: Mr. John Underwood, Center for Naval Analyses
(18) *Techniques for Modeling Tactical Nuclear Warfare*: Mr. Stanley Spaulding, Vector Research Inc.
(19) *Logistic Support and Combat Unit Effectiveness*: Mr. Ellwood Hurford, U.S. Army Logistics Center
(20) *Comparison of Results from IDAGAM with an Aggregated Combat Model*: Mr. Norig Asbed, Army Concepts Analysis Agency
(21) *Modeling of Tactical Decision Processes in Theater-Level Gaming*: Mr. Robert Robinson, Headquarters U.S. Air Force, Studies and Analysis

Wednesday, 28 September, Afternoon Schedule

DISCUSSION PANEL
Mr. Robert Farrell
Mr. James Dunnigan, Simulation Publications, Inc.
Dr. Robert Helmbold, R&D Associates
Dr. Klaus Niemeyer, IABG, Federal Republic of Germany
Mr. Standlee Steenrod, Army Concepts Analysis Agency
Prof. James Taylor, Naval Postgraduate School

SESSION IV: Game Theory in Theater-Level Modeling: Optimal Solutions and Heuristic Solutions
Chairman: Prof. William Lucas, Cornell University
Opening Remarks: Prof. William Lucas
(22) *The Role of Game Theory in Force Planning and Game Associated Problems*: Prof. Martin Shubik, Yale University
Solution Procedures for Multistage Games at the Theater Level
(23) *DYGAM: An Algorithm for Solving Multistage Games*: Dr. Zachary Lansdowne, Control Analysis Corp.
(24) *Descriptions of the Solution Procedure Used in ATACM*: Dr. Frederick Miercort, Consultant

Thursday, 29 September, Morning Schedule

SESSION IV (Continued)
Man/Machine Interactive Gaming and Heuristic Solutions
(25) *Game Theory in Theater-Level Modeling: Optimal Solutions and Heuristic Solutions*: Dr. Alexander Dobieski, TRW
(26) *Some Common Problems with Man/Machine Interactive Modeling*: Dr. Paul Tuan, SRI International
(27) *Problems in the Applications of Game Theory at the Theater-Level*: Prof. John Mayberry, Brock University

DISCUSSION PANEL
Prof. William Lucas
Prof. James Falk, George Washington University
Dr. Lola Goheen, SRI International
Dr. Jeffrey Grottle, Institute for Defense Analyses
Dr. John Tomlin, Institute for Advanced Computation

SESSION V: Data Base Requirements for Theater-Level Models and Data Related Input Generation Problems
Chairman: Mr. Robert Schneider, Office of the Secretary of Defense, Planning and Evaluation
Opening Remarks: Mr. Robert Schneider
(28) *Overview of Data Base Problems in Theater-Level Modeling*: MAJ Raymond Bednarsky USMC, Command and Control Technical Center

Thursday, 29 September, Afternoon Schedule

(28) *Overview of Data Base Problems* (Continued): MAJ Raymond Bednarsky USMC

DISCUSSION PANEL
Mr. Robert Schneider
MAJ Brian McEnany USA, Joint Chiefs of Staff (Studies Analysis and Gaming Agency)
Dr. Jacob Stockfisch, American Petroleum Institute
LTC. Donald Berg USA, Joint Chiefs of Staff (Studies Analysis and Gaming Agency)

FINAL SUMMARY SESSION
Dr. Frank Kapper, Session I
Dr. Jerome Bracken, Session II
Mr. Robert Farrell, Session III
Prof. William Lucas, Session IV
Mr. Robert Schneider, Session V

List of Attendees

Name	Affiliation
Anderson, Bruce	Institute for Defense Analyses
Asbed, Norig G.	Army Concepts Analysis Agency
Bednarsky, Raymond MAJ USMC	Department of Defense Command and Control Technical Center
Berg, Donald LTC USAF	Joint Chiefs of Staff (Studies Analysis and Gaming Agency)
Bode, John	The BDM Corp.
Bonder, Seth	Vector Research, Inc.
Bracken, Jerome	Institute for Defense Analyses
Bruni, Jerome CAPT USAF	Defense Nuclear Agency/Kirtland AFB
Dare, David	Defence Operations Analysis Establishment (Great Britain)
Dargis, Gerald P.	Central Intelligence Agency (OSR)
Dobieski, Alexander	TRW Defense and Space Systems
Doty, Robert LTC USAF	Commander in Chief Pacific (J021)
Dunnigan, James	Simulation Publications, Inc.
DuPuy, Trevor COL USA (Ret.)	Historical Evaluation and Research Organization
Falk, James E.	George Washington University
Farrell, Robert	Vector Research, Inc.
Felch, John E., Jr. MAJ USA	U.S. Army War College
Ford, Lester	General Research Corp.
Glassman, Neal D.	Office of Naval Research (Code 434)
Goad, Rex	SHAPE Technical Centre
Goheen, Lola	SRI International
Grotte, Jeffrey	Institute for Defense Analyses
Grace, J. J. COL USMC	Headquarters U.S. Marine Corps.
Helmbold, Robert L.	R&D Associates
Hess, Carl H. COL USA	Office of the Assistant Secretary of Defense (Material Acquisition and Logistics)
Hetherington, Jay A.	Central Intelligence Agency (OSR)
Hoeber, Frank	Consultant
Huber, Reiner K.	Hochschule der Bundeswehr München, Federal Republic of Germany
Hudson, John R.	Canadian Department of National Defense, Operations Research Establishment
Hurford, Ellwood	U.S. Army Logistics Center
James, Ronald CDR USN	Office of Naval Research (Code 431)
Jayachandran, Toke	Naval Postgraduate School
Kapper, Frank	Joint Chiefs of Staff (Studies Analysis and Gaming Agency)
Karber, Philip	The BDM Corporation
Karr, Alan	Johns Hopkins University
Kent, Glenn LGEN USAF (Ret.)	Glenn A. Kent Associates
Kerlin, Edward	Institute for Defense Analyses
Konkel, Dennis	Department of Defense, Command and Control Technical Center
Kovanic, Frank	SRI International
Kugler, Richard	Office of the Secretary of Defense (Planning, Analysis and Evaluation)
Lansdowne, Zachary	Control Analysis Corp.
Linnenkamp, Hilmar	Ministry of Defense, Federal Republic of Germany
Louer, Philip	Department of the Army, Office of Deputy Chief of Staff for Operations and Plans
Low, Lawrence	SRI International
Lucas, William	Cornell University
Lundegard, Robert J.	Office of Naval Research (Technical Director)
March, Miles COL USA	Arms Control and Disarmament Agency
Marshall, Andrew	Office of the Secretary of Defense, Net Assessment

Name	Affiliation
Mayberry, John	Brock University, Ontario, Canada
McEnany, Brian MAJ USA	Joint Chiefs of Staff (Studies Analysis and Gaming Agency)
Miercort, Frederic	Consultant
Minckler, Rex	General Electric, TEMPO
Niemeyer, Klaus	Industrieanlagen-Betriebsgesellschaft mbH, Federal Republic of Germany
Normann, Soren	Boeing Computer Services
Olsen, Paul	Institute for Defense Analyses
Osterman, Gene	Lulejian
Pappas, Aris A.	Central Intelligence Agency (OSR)
Pugh, George	Decision Science Applications, Inc.
Redelman, Roger LTC USMC	Joint Chiefs of Staff (Studies Analysis and Gaming Agency)
Robinson, Robert	Headquarters, U.S. Air Force, Office of Assistant Chief of Staff, Studies and Analysis
Schneider, Robert	Office of the Secretary of Defense (Planning, Analysis and Evaluation)
Shewmaker, John	Congressional Budget Office
Shubik, Martin	Yale University
Sievert, Volkmar	Ministry of Defense, Federal Republic of Germany
Simpson, J. Randolph	Office of Naval Research (Code 431)
Smith, James R.	Office of Naval Research (Code 431)
Sovereign, Michael	Naval Postgraduate School
Spaulding, Stanley	Vector Research, Inc.
Starkey, James	Central Intelligence Agency (OSR)
Steenrod, Standlee	Army Concepts Analysis Agency
Stockfisch, Jacob	American Petroleum Institute
Strudeman, Richard COL USA	Joint Chiefs of Staff (Studies Analysis and Gaming Agency)
Taylor, James	Naval Postgraduate School
Thackeray, Ross	Ketron, Inc.
Tomlin, John	Institute for Advanced Computation
Tuan, Paul	SRI International
Underwood, John	Center for Naval Analyses
Varley, Thomas C.	Office of Naval Research (Code 434)
Walker, Lanny LTC USA	Third Basic Combat Training Center, Fort Dix
Wegner, Louis	The Rand Corporation
Welch, Jasper, Jr. MGEN USAF	Headquarters U.S. Air Force, Assistant Chief of Staff, Studies and Analysis
Young, G. C. Saul, Jr. MAJ USAF	Air Force Institute of Technology

Proceedings

Chairman — Opening Remarks

LAWRENCE J. LOW
SRI

Mr. Low: Gentlemen, I certainly welcome you all here to Xanadu, and it gives me great pleasure to be up here this morning before a distinguished audience to call this meeting to order. I do hope that you'll find this meeting to be stimulating and even enjoyable perhaps, and that it will be nonetheless useful and productive in allowing us to identify and to define some key problems in the complex field of gaming that are in need of attention.

The meeting has been structured in a way that undoubtedly sacrifices some depth for breadth in the areas to be covered. This has been done to afford the recognition of interfaces between the areas of simulation, optimization, and data base development, so that such things as trade-offs or the employment of recursive techniques between these areas can be more readily identified and examined. In addition, the consideration of user needs and the links between these needs and any of the analytical techniques one might employ is another matter that requires serious attention, as does the ever present problem of promoting better communications and understanding between the research, the analysis, and the user communities.

Note also that we're stressing "theater level" and "force planning" in the meeting. This is very important, really. We have to do this to define boundaries for our discussions since there are literally an infinite variety of games and types of problems to which they can be applied.

I'm sure we'll find ourselves drifting from the "straight and narrow" scope definition from time to time in our discussions, and I would say that some of the drifting would be unavoidable and, in fact, may be quite proper since you can't discuss the upper levels of a hierarchical phenomenon such as gaming without bringing in some of the lower supporting levels. Let's try, however, not to get too sidetracked in our discussions, or worse yet, derailed. And please do remember that it's the use of gaming as a planning device rather than as a predictive device that we're concerned with here.

I would hope that we can give precedence to substance over form at the workshop, and in this connection I think the meeting should be as informal as we can afford to make it, while still permitting us to cover the material spelled out in the agenda. Above all, let's have candor prevail in the papers that are presented and the discussions that follow.

Well, without further ado, then, I'd like to turn the podium over to Dr. Varley of ONR who has a few remarks to make and who will introduce the keynote speaker, General Welch.

Sponsor's Welcome

DR. THOMAS VARLEY
ONR

Dr. Varley: Good morning. I'd like to welcome you here on behalf of Admiral Geiger, the Chief of Naval Research and also on the part of the workshop steering committee who put this together. The Office of Naval Research is the primary research component for the Department of Navy and is always looking towards some fundamental truths to extend our understanding of the problems within the Navy community with the objective of solving those problems and thereby increasing the efficiency and the effectiveness of Navy operations.

Workshops are one of the ways that we do this. We attempt to bring experts together from different fields to discuss the problems, the approaches, the state of the arts, the pitfalls, identification of the voids that we have in our knowledge, and then to take that, put it together, understand where we are, where these voids are, and then how we can use this information to move forward with a research plan to solve these problems. That's the kind of a thing that we hope to get out of this gathering, and, as a workshop — again emphasizing something that Larry said — we do mean work. It's a give and take operation, constructive criticism is welcome and expected. We expect each of you to give your opinion when you feel justified to do so, so keep it very casual and do speak when you feel that there's something that needs to be clarified.

Now then, to our keynote speaker for today. I have mentioned experts — each of us considers ourselves experts in different kinds of areas, and I want to give you just a little bit of background on our expert today. Number one, he's a Ph.D. in physics from the University of California at Berkeley, a distinguished graduate of the Industrial College of the Armed Forces, has had duty at Lawrence Livermore Laboratory, has been a science advisor for the Air Force on special weapons design, has had duty at Rand, has had duty in the Air Force and at OSD in systems analysis, has served as Special Assistant to the Secretary of Defense for Atomic Energy, and is currently the Assistant Chief of Staff for Studies and Analysis, Headquarters, U.S. Air Force. He's a scientist, he's a military officer, he's a planner, and he's an analyst.

It's a great pleasure to introduce Major General Jasper Welch, Jr., U.S. Air Force.

Use of Gaming in Force Planning

MGEN JASPER WELCH
HQUSAF (AFSA)

Cutting edge issues and problems in the development and use of hierarchical models and large-scale simulations by the Air Force.

I'd like to talk about three different aspects — that is, put on three different hats. I'd like to talk as a user; that is, as a consumer of the output of studies using theater level analysis; second, as a manager of a group of working analysts, which I inherited from Glenn Kent and our mutual good fried Bob Lukeman; and, third, as a sometime working analyst myself, and tell you a little bit about our current "edge of concern" problems in studies and analysis.

I think that from my point of view we've *had* to go to the theater level of gaming, and I say it that way because in the old days life was simpler, for a variety of reasons. I think that you may remember that the earliest applications of force planning was the quantitative assessment of alternative systems. Now, when we were comparing, for example, the Air Force Thor and the Army Jupiter systems and some of the heyday items of the interservice competition in the late 1950s, that became almost like service source selection, but as we went on, things got a little bit more complicated, and I guess the real crunch was when a thing called the high-low mix was invented. Now for whatever reasons you may have thought that that was invented, it nonetheless brought you face to face with a problem where the proposition became mixes of more sophisticated and expensive platforms in fewer numbers and larger numbers of less expensive and less capable one-on-one platforms. There are many, many examples of that. The Air Force example is the F-15 and F-16 mix, and you can sort of tell which is the high and which is the low because one of them has two engines and two tails and the other one has one engine and one tail, and the engines are exactly the same as a matter of fact. It gets a little bit more difficult to tell sometimes, but, generally speaking, the high end of the mix costs more.

Now, that brings you face to face with trying to evaluate these things in a milieu of other equipment because the less capable airplane by and large depends more on the supporting infrastructure of information systems, radar systems, some difference sometimes in supportability, different ranges sometimes, and, therefore, different base structure sometimes. Therefore one structure simply will not reflect the essence of the high-low mix. Not that you will make a numerical mistake, but you won't even get the point of the high-low mix unless you have the interactions of the individual aircraft with its maintenance infrastructure, information infrastructure, and command infrastructure.

Now, there are some other things which have driven us to theater level modeling. I think that in the Air Force one of the things has been the emphasis that the Soviet Union has placed on surface-based air defenses. As long as the main threat to airplanes completing their mission was duelling in almost a medieval fashion with other airplanes in dogfights, à la the Red Baron, then indeed, models of airplane-on-airplane (or at most a "few-on-few") were adequate to sort of get the exchange ratio and the overall engagement rate straight, these being the two major factors to understand in opposing systems. The exchange ratio is how many of Red to how many of Blue, and the engagement rate is how many of either per unit time as measured by a clock tied to other events — for example, the land battle, or the state of negotiations.

Now, you add surface-to-air missile systems, and furthermore you find the Soviets doing things like overlaying, as we've discovered in East Germany, five separate air defense systems, and I mean systems in not the weapon system SA6 type, but five separate and distinct total infrastructures. There's the East German air interceptor force, and the old SA2, with all of its GCI radars. Overlaid on that is an SA3 system, and then they give the SA2's to the East Germans, and then there are the 4s, 6s, 8s, 9s, and 7s and all that stuff that moves with the Army, so there are five separate overlapping systems all available in the same air space at the same time and having different command and control linkages, different functions, indeed even different political functions, so that that type of synergism between the opposing air defenses and the concommitant necessity to have a multilayered approach to mitigating that attrition leads you inevitably into larger than one-on-one situations.

You can give a similar type of long and complicated example having to do with precision-guided and stand-off munitions. There's another thing that has come into the question which is related to the high-low

mix; that is the overall force mix question. The Air Force is currently in the situation of having invented, designed, and having in production four or five new aircraft types. What is at issue is at what rate shall we modernize? That is what should be the total number of new aircraft per year. What shall be the mix of the modernization program? And, to some extent, to what ultimate force level are we building building?

Now, all of those things clearly require not only theater-level modeling of something like the European situation, but something which is quite beyond our current capability, namely to say, well, while you're fighting on the central front in Europe, to what extent can you afford to draw down on forces in other parts of the world — out in the Far East and the Middle East and North Norway and so forth. It turns out that fully a third of the Air Force is sort of generally contemplated to be deployed in those areas even in the event of substantially tough fighting on the central front.

Now, we also have another thing, and, in fact, if I remember right, this is one of the first successful applications of game theory to changing people's minds about large-scale force decisions, and that is the trade-off between airplanes and air bases. Back in 1970-1971, and I can't quite remember the thing, when I was working for Glenn Kent and young Goodson was working with us, we did a thing which many of you know about here, primarily from its theoretical interest, called TAC CONTENDER. That was a game theory approach on a very, very simple sort of level, but it seemed to us to be adequate to encourage people to make trade-offs in a way between shelters on airbases and more airplanes, and, indeed, the historical track is that the analysis did in fact swing a number of influential people in the Allied air forces to support an American initiative for making shelters on airbases qualify for NATO infrastructure funding. Without that qualification the U.S. Congress won't put any money in nor will anybody else, and so that really made a difference. Again, theater-level models are necessary to make those kind of trade-offs between airplanes and airbases.

Now, the second question is, as a manager, how to go about doing theater level gaming, and I might say that there are sort of two general ways to go about it, and we've done both. One way is to make one big "humongous" model, modular to be sure, and all of those sorts of things, but nonetheless, one big "humongous" model, and we did that — it's called the Advanced Penetration Model. It has been running for some number of years, and it only took a finite amount of time to develop, but it was a fairly large finite amount of time. During that period of time, first Glenn and then Bob, and to some extent myself, had to fend off people who said, "When are you ever going to pay the rent; when are you ever going to do anything other than just develop that model?" And, indeed, there's a sort of a problem because I am still now paying of the order of half a dozen people in-house and an equivalent number of technical contractor support simply to maintain the model in a happy state by somebody's definition.

Now, what I want to do is distinguish between that approach and something else that we're now trying, which is really what I want to talk about this morning — that is hierarchical models where the size of the model is sort of independent of its level — that is, wehre you might have a theater model which is so big in numbers of cards, and then there is one that's a little bit lower that takes care of some part of the operation, which is only loosely coupled to another. For example, one way to think about it is to look at the Soviet plan and see that they're talking about armor breakthroughs, and maybe you want a model which will handle a breakthrough, recognizing that each breakthrough is only loosely coupled to the adjoining breakthough; at least that's the way the guy on the offensive has in his mind how the war's going to go. And, again, maybe you can just run an airbase and solve to that kind of a model, not the least of which is a very practical one, which is that I can organize my activity, my people, exactly the way the model is organized, so that I can have a group working on the top part and then three people working on the three down on a lower tier and five working on the next tier below. It's not quite that way, but at any rate I can have tiers, with each of these models at different levels. If I organize it furthermore so that it is not only a hierarchy with regard to the logic of coming to a theater model, but it also has outputs which are relevant to real decisions that real people make, then I can pay the rent because I can use these intermediate level models and lower levels to actually lend some good advice, or bad advice as the case may be, toward actual problems being asked of me by real decision makers on a day-to-day or week-to-week basis. We have by and large done that, somewhat by accident, but when we recognize the value of this approach to decision-making at other than theater-level, we have now tended to encourage that sort of thing.

Now, having told you about why it's such a great idea, let me tell you about some of the problems that we're finding, and leave you with that.

The first problem is that, clearly, if you're going to have a hierarchy of models, they have to have some amount of interoperability. Now, I'd like to sort of focus on that word and use it as a shorthand for a variety of fairly ill-defined worries and constructs throughout the rest of the talk. So — you have to have some sort of interoperability, if you're going to have this hierarchy that says the one-on-one models have to feed the many-on-many which have to feed the theater level. Within that general scheme of interoperability, probably the most difficult problem is that of consistency. That is, we have an item — and I've used it this morning

and you all nodded — the F-15. Now, that's a code in the perfect sense of a semantic code. An F-15 evokes in your mind all kinds of things depending on your background, your problem set, and what you're trying to do. Now, we just finished having a great big fight with the OSD analysts because what the code F-15 evokes in their mind is noticeably, distinctly, and to our point of view, deleteriously different from the notions that rise in our minds in regard to the code F-15. In particular, there are honest uncertainties, I won't say differences of opinion; it's even more fundamental than that; there are honest uncertainties with regard to certain technical performance parameters of the F-15 and its associated armament, and that's no great secret. There was a hellacious fight about it associated with the first statement that Harold Brown made to the Congress this session. So there really *are* those sort of inconsistencies. So when people say, "I have run a program," and they have a diagram, and a nice little boiler-plate box up there like Glenn and I taught people to do, and it says F-15, you still don't know if it's the same F-15 that somebody else had in mind. You don't know if they used the same P_ks; you don't know — at the grossest level and way down into the subtle guts of the thing — that a rose is a rose is an F-15. So, you really have to worry about that.

Actually, we don't have any substantive literature to which a new analyst can go and say in a general way, "How can I check for consistency?" We don't have a set of substantive literature such as that in which you can find out what it means if the noise of a parametric amplifier is so many dbm, or dbw. You can't find that out. There are no standards. So that's really a problem.

Now, there is a problem with another thing — something I call enforceability. In a one-on-one set of engagements we make certain assumptions about how things start. In an air-to-air engagement, how does the battle start? The guys are way apart and have no contact with each other but are generally flying in the same general area. You have to know how the battle starts or else you'll get nothing for having spent a lot of computer time. So you generally bring the guy close enough that he has half a chance of interacting. That's the most efficient way, right? Now, the question is, when you embed that engagement in the hierarchical model, how do you enforce the fact that the aircraft had half a chance to engage? Have you really paid the price in some other area, in some command and control structure or whatever, to indeed give the guy half a chance? Well, let me tell you why it's important. You see, it could turn out that if the engagement rate is such that the guy gets half a chance then the exchange ratio may turn out to be three to one, that is, Reds versus Blues. On the other hand, if the engagement rate is reduced, either by accident or on purpose, (the extent to which one wants to do it by accident or on purpose depends on how your exchange ratio came out, right?) — if the engagement rate is reduced, then the exchange ratio may be widely different. So you have to worry about the enforceability of the assumptions on the lower level hierarchy. The question is — How do you enforce those assumptions to be valid when you use the results of the low level in the high level?

Related to that is a question of initial conditions and boundary conditions. Again, in our one-on-one engagements, in trying to get our air-to-air data into TAC WARRIOR, which is currently our best and brightest theater level model, we found that the one-on-one air people have been trying to solve different kinds of problems. What they've been interested in was whether they should add another fuel tank to the F-15 or not, should they have an A-9L missile or an A-9J plus missile. They were looking at design modifications to the F-15 plus its armament in order to decide what combination of things to buy when they decided to buy the code F-15. Now, one of the things that happened then was that they looked at the F-15 under a variety of combat conditions and as a particular item, which, when it was pointed out actually the guys were a little embarrassed about it. They had allowed the dogfights, the one-on-one dogfights to continue down to really low levels, you know, they drove them into the ground almost literally. The airplanes would continue to lose energy and yet the investigation of the process would continue on down in an attempt to see what would happen. The result was the F-15 wound up beating Fishbeds all the time, but the F-15 was at 200 feet and 200 knots and was sort of, you know, available for easy plucking from any other airplane just coming by. So if you went back to more realistic cut-off conditions, recognizing that there were going to be other airplanes in the area, conceptually at least in the theater model, then what you found out is that the engagement rate suffered a little bit but the exchange ratio suffered quite a bit, that is, he got a fleeting shot at the guy and then he had to recover and go back, but he didn't use as much fuel so he got another chance to do it. Well, it's these types of interfaces that are really quite important.

The next thing I'd like to talk about is the question of sensitivity analysis. Here, hierarchical models actually have some significant benefit. I once gave a talk at the ORSA/TIMS meeting, which was sort of entitled "Ten to the Thirtieth is Forever." Let me just give you a couple of numbers so you get a sense of what that means. Consider a model that has 100 inputs, and we want to do a sensitivity analysis of the output with regard to the inputs — 1 output, 100 inputs. Now, 100 inputs isn't all that many. We've got lots of models with lots more than that, but our minimum sensitivity would be to take two values for each input, which is 2^{100} runs, or 10^{30} runs. Now suppose that we have a very efficient computer such that for a given set of inputs we get the output, and the processing time is one nanosecond. Okay, now 10^{-9} times 10^{30}, so that's 10^{21}

seconds, which is 10^{60} days, is 10^{13} years, which is 10^{11} centuries. That really is forever, and so clearly, you can't do sensitivity in that sort of crass sense, with anything like the computational power at hand today. Now, interestingly enough, to follow that same analysis through you'll find that, in fact, you can do it with about 20 inputs quite handily and you won't get thrown off the computer. Do you know how long you can run one run? A million seconds. That's all you can do on any computer anywhere in the world because that computer is never owned by a person, it's always owned by an institution and there's no institution that will let you run more than a million seconds. A million seconds is about ten days.

Now, if you can group inputs in a hierarchical way then, of course, you can substitute things. That is, you can just use the equivalent radar range and vary that, and, of course, there are many ways to get that radar range the way it is. That's what you mean by a real hierarchy, and, remember, people really do design radars, and there really are more than 100 inputs, and they really do think they have done a fair amount of optimization because there's a lot of cost competition in that business these days. So you know it's possible. You really have to think through the hierarchical structure because you aren't going to know the inputs well enough, and if you're in the business of influencing decision, you had better be able to tell a guy how likely you are to be off the mark. That's just a general rule.

The next thing I'd like to talk about is preprocessors. Now, we have, in studies and analysis, a matched set of very detailed surface-to-air missile models. That is where the 4s and the 6s and the 8s and the 9s are all on the same basis. And that's very helpful because if you go to the intelligence people they'll tell you this division has six of those and five of those and three of these. So you don't want the different kinds of SAMS to be unmatched in their one-on-one characteristics because you'd like to balance your tactics against the mix that you're going to find. So we have this matched set, and they're fairly detailed, you know, they've got fifth order Runge-Kutta and all that stuff. And you ask the guy, why all that, and he can give you a very good reason for almost all of it, at least he can wear me down. But, I notice a curious thing when I look at the output and that is, most of the time, the output is predictable to within two numbers. That is, it's either one for the P_k or zero. Ninety-nine percent of the outputs are either one or zero. Now, if you plot the answers on an array it turns out you can do cross correlation coefficients. If there's a one in a certain place in the array, the adjoining numbers are more likely to be ones than zeros, and, in fact, you check for that to see if you think, the model is working right. So that, when it's an easy shot, the guy is essentially reliability limited on the P_k, and, when it's a really tough shot, it doesn't matter if it's tough for three reasons (any one good reason will do). There's a fairly narrow boundary at the edge of the operating envelope where the missile just can't quite make the corner, or where the target is just out of range, or the cross section is just too low, or the target is down in the clutter just too far, or there's a little feather edge, or a big feather edge but generally it's a little feather edge. Now, at the same time you build the one-on-one, one ought to be made to make a preprocessor that will announce with fair fidelity (but even sloppy fidelity will do), that I'm in the high P_k region, the essentially zero P_k region, or probably I'm in the feather edge. Then, if we do that, the preprocessor can be used, as you can clearly see, in a hierarchy to save a lot of time, a lot of effort, and, indeed, to give you greater insight. This is because one of the things that the preprocessor is bound to do while it is operating is to announce when and why it is in the zero region or why it thinks it is in the zero region. And that is generally the clue to the answer the decision maker asked you in the first place, which is, "What am I supposed to do about this problem?" He's not really interested in the P_k. You may be interested in the P_k. You made that up. What he's interested in is what should he do. And it may be that the preprocessor is great for finding that out. Now, you want a high fidelity preprocessor just to confirm things and make sure that you didn't get zero because of a logic fluke in the FORTRAN statements. By and large you should use a preprocessor. It is not routinely done, and, if I'd have known how useful it was, I'd have started trying to do it long ago.

Okay, one more piece of pseudoscience. We had a thought to relate it to these two big models that we have running. We have a theater model of bomber penetration, which we call the Advanced Penetration Model. The structure of model allows us to follow every bomber's tail number, and associated with that bomber's tail number is what we call a state vector, and there are many many entries in the state vector, many elements to that vector, some of which are its position, its fuel, state of armament, its current velocity, its current altitude, and whether it's engaged. The process of the computation is to event step and to update the state vectors of all the airplanes, move then forward, run them into SAMS, etc. and we don't keep track of anything else. We keep track of where the airplanes are and we have state vectors for the SAMS and all that sort of thing so the focus is on empties. Now, in TAC WARRIOR, which is our big theater model, it is done quite differently. There, we have boxes, and we keep track of how many airplanes are in the boxes, and the boxes are described in terms of where, when, and what you are doing. That is, there's a box that says how many airplanes in this airbase are in unscheduled maintenance, and when they get through with unscheduled maintenance they move over to refueling, and when they get through with that they move into a hangarette

and they sit in the hangarette and then they're called on to go out and they sit on the taxi-way and then they take off. For every one of those little way stations, as though those planes were going down a production line, we keep track of how many are there, and the rate at which they move through. This rate depends on how many avionics technicians, etc., you have, and then when they get up in the air, the rate at which they get shot down depends not only on how many are there but on how many other fighters are in that same general area. There are planes here, and planes there, meaning that they'll fight, and some of those will get shot down. Well, we noticed that that is analogous to hydrodynamics in which there are two representations of the flow. One is called Lagrangian, the other is called Eulerian. Now, the Lagrangian flow is when you sit in a boat without a paddle and float down the river and observe what happens; you move with the moving fluid. The Eulerian is when you sit on the river bank sipping a mint julep and watch the water go by. Advanced Penetration Model is Lagrangian. We sit with the entities and follow them around and notice what adventures they get into. In the Eulerian, we sit at the maintenance hangar or wherever and know that the airplane's coming by. Now, we didn't think about it this way at the time we did it, but I think that there are natural reasons for doing that. In the strategic bomber problem, there are thousands of miles between airplanes, certainly hundreds of miles, and why keep track of all that empty space? Nothing's happening there. On the other hand, in the central front, there are airplanes everywhere, and I don't really care about one F-4 versus another F-4. All F-4's look alike, but I've got these big production factors I've got to worry about to crank out sorties and so forth, and it's a very overfilled state. That is, for every set of descriptors I can list, there are many airplanes with equivalent descriptions at the same time, and it's natural that it's an overfilled state space, so I use the Eulerian representation.

Well, I've never seen it described that way, but I thought you might find it interesting. It could very well be that a proper hierarchy will start Langrangian and transform itself to be Eulerian as you get the higher and higher levels of aggregation, and if you don't see that happening I suspect you're going to be pretty mixed up half way through as your frame of reference changes and you don't know that it is happening.

All of that's by way of saying that I think we need a tremendous number of theoretical constructions. I'm interested in this for two reasons; one, I'm just sort of interested, like you'd be interested in mathematical puzzles; second, I think we have all been just terrified at the number of bad scenes that we have perpetrated on the taxpayer, if no one else — of starting down paths and really stumbling very very badly, and never getting anything worth shouting about. It seems to me that a lot of that has to do with the fact that we don't have adequate theoretical construction to describe what it is we're doing and to know in a fundamental way where we're headed.

I heard for the first time what I now find to be a particularly old idea, which is, if you don't know where you're going almost any path will get you there.

As you know, there is a great deal of interest in injecting a higher degree, or some degree, of overall management into the studies analysis world of the DoD. I am not particularly keen on that idea. I have objected to many of the attempts to do that. I will continue to object, but we won't be successful until we show that as a profession we can run our own affairs in a professional and sound way. It's a peculiar profession in the sense that there is a very limited commercial sector, and most of us are engaged in activities which are funded by the taxpayers treasury and, therefore, we have a special responsibility to spend that money wisely. We really have to do better. It's very easy to say that, and a lot harder to do it, but I don't see how we can avoid some pretty onerous problems unless we get a much tighter handle on how we're proceeding. Thank you very much.

Session I — Gaming Utility from the User's Viewpoint

Opening Remarks

DR. FRANK KAPPER
JCS/SAGA

Dr. Kapper: I would like to talk a little bit about the objectives of this session, a little bit about the focus, and a little bit about what the modeling community is.

The objectives are to identify and discuss the types and decisions at various governmental levels that are supported by computer gaming. Second, to identify the more serious deficiences and shortcomings as the user sees them in current practice and techniques.

While the focus is on theater-level gaming and specifically in the formulation of plans and policy for this particular session, I think that some of the speakers will refer to other than just theater-level war gaming because there are many examples which speak to the objectives and theater level gaming itself may be in some sense an artificial constraint.

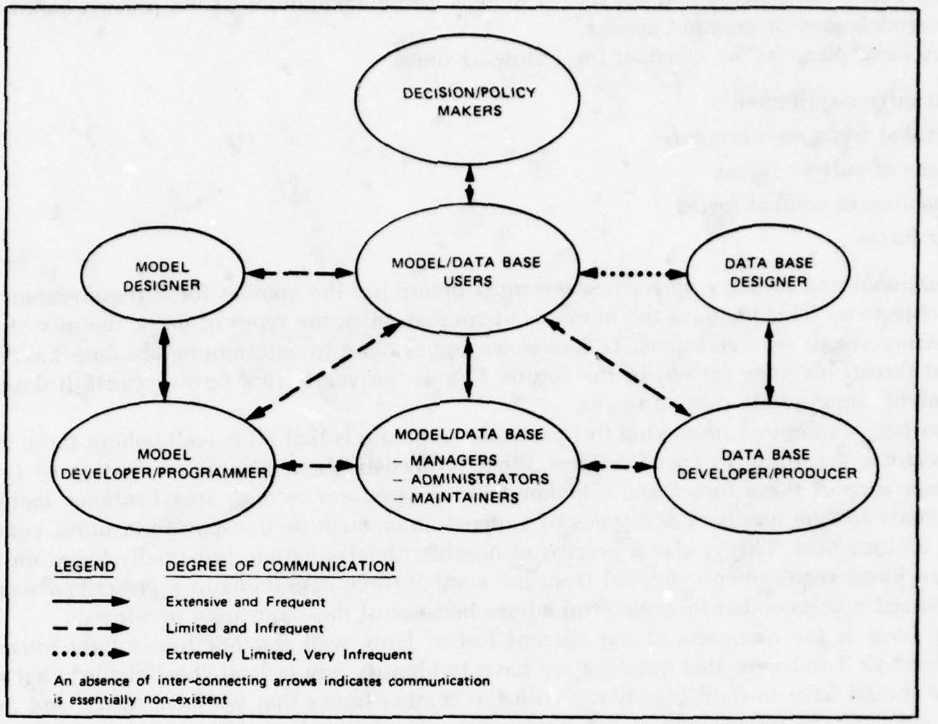
Now, I'd like to talk just a second about the users' perspective. Why would a user be interested in models? Well, you say, that's pretty simple, either they help you do your job as an analyst, as a policy maker, or an operator, faster, cheaper, easier, or better, or else they're not really worth the investment. So, it's kind of a simplistic approach, but, in essence, if you just concentrate on those terms, that's really why we are in the business. Otherwise, as I say, I don't think we should use models.

Now, who makes up the modeling community? Obviously, the users do, but the same can be said for the data base and the model administrators, operators, and maintainers. The models designers and developers are also included, and so are the data base providers. In most cases we're really talking here about the intelligence community. (See Slide A)

One of the problems that I think we have as users, and I talk about it in a very general sense, is getting the decisions makers, the policy makers, involved. Now, those I see here, with notable exceptions, are principally analysts, modelers, and model users, if you will in that sense, and administrators, maintainers, data base collectors, and managers. I see people who influence decision makers, but I see very few decision makers here. Perhaps that poses a question for the future, when we leave this symposium. The essence of that question is the communication back to, and with, that intersecting community of policy and decision makers. I think we should ask ourselves throughout the symposium the question "how well and often do we in the modeling community, communicate with each other?" Do we really know what the decision maker needs, the policy maker? Do we take a long-range perspective relative to his mission, to his function, to his problems, or are we essentially job shoppers waiting for the next task to come in so that we can use our tools on it?

Anyway, with these thoughts I'd like to introduce Mr. Phil Louer, Department of the Army, Office of the Deputy Chief of Staff for Operations and Plans. He's been in the modeling business so long that he'll be able to touch on many of the things that are of interest to all of us.

SLIDE A RELATIVE DEGREE OF COMMUNICATION BETWEEN USERS, DESIGNERS, DEVELOPERS, PRODUCERS AND MANAGERS OF MODELS AND DATA BASES AND DECISION/POLICY MAKERS



1 — Theater-Level Gaming in the Formulation of Plans and Policies (Army)

MR. PHILIP LOUER
Army DC/S Operations and Plans

The Army Force planning process and the supporting role of theater models.

Mr. Louer: When Larry (Low) called me a few weeks ago and asked me what I thought of the preliminary plans for this workshop, I kind of hesitated for a moment, and then I decided to give him an honest answer and I said, "Well, it looked to me like it was pretty heavily oriented toward the theorists' point of view and not enough from the users point of view." So the next question he asked me was, "Well, how about you giving the users' point of view from the Army," so I'm here to do that. I feel somewhat like a minister about to preach a sermon on righteous living, but I believe in what I'm going to say, and I hope I can get the point across.

I'm going to talk about theater model use in the Army force planning process and how applications of theater models can develop guidance. Through discussions of the kinds of knowledge to be gained from these applications, I'll then lead into some of the characteristics that are required in theater models. Finally, I'll discuss briefly the Army's use of theater models in the force planning process.

As I talk about this process and the use of models, some adaptations of the process based on the limitations of the models may or may not appear.

In Army force planning we consider the following items:

- Combat force requirements
- Non-combat force requirements
- Readiness of current forces
- Sustainability of combat forces
- Nuclear forces

Given our national and military objectives, we must determine the combat force requirements necessary to meet these objectives. This includes the numbers of combat units, the types of units, the mix of combat arms, and active Army versus reserve forces. This exercise comes close to determining absolute numbers to oppose an estimated threat. It's done far out in the future, five or ten years from now, because it does take time to acquire or modify appropriate combat forces.

An important assumption in making this particular estimate is that all is well behind these combat forces. They can perform according to specifications. What is needed, then, is a determination of the noncombat forces that can support these forces and whether they can perform to their specifications. Included here are the requirements for the numbers and types of support units, such as transportation units, communications, intelligence, and medical. This is also a process of absolute determination, principally based on the organizational and workload requirements derived from the combat force operations in a prescribed scenario. This is also synchronized to the combat force determinations because of the acquisition problems.

Another issue is the readiness of our current forces. How well can our forces fight today if it should become necessary? To answer this question we have to identify any critical shortfalls in what we have versus what we should have to fight effectively with the combat forces that we have today, and in this process we have to measure the importance of these particular shortfalls. Now, we're addressing shortfalls in such areas as personnel, material, weapons, ammunition POL, deployment capabilities, and mobilization capabilities.

Our objective in this analysis is to determine the seriousness of the shortfalls that do exist, and to establish priorities for improvement over the upcoming years. Now, eventually, this information is used to guide fund allocations for improvement of readiness in our forces.

For sustainability, we need to know what stocks should be kept on hand to support a war. This is a determination somewhere between two limits. As a minimum we should have enough so that we would not lose the war because of inadequate supplies. On the other hand, we don't want to require so much that we cannot procure it and store it.

The nuclear issue is a contrast from the above process. Here, the emphasis is on individual weapons, launching systems, and the particular forces associated with guarding, handling, and delivering the weapons. Other planning issues are also addressed, however. These include nuclear employment doctrine, political and military constraint effects, the security of the weapons in storage and handling, thresholds for a nuclear war, and effects on a conventional war of the presence, or use, of nuclear weapons.

Using Slide 1-1 I'll discuss certain characteristics of the models which are required to address the particular readiness issues I outlined above. Combat forces must be represented in theater models by identified units through the command organization, and this means all units, all command and control, and force disposition. The forces may not be resolved necessarily to battalion level, but they must reflect the battalion mix of combat arms and associated interactions. Of course, the capabilities of the weapons present must be reflected, as well as fire support elements, the artillery, attack helicopters, and close air support, and the combat engineer contributions to combat effectiveness.

With regard to the noncombat forces, the effects that these forces have on the combat must be represented along with any results of combat that would degrade the capability of these forces to support combat forces. As a minimum, the functions of communications, intelligence, logistic, and medical units should be there.

The information flow, knowledge of the forces, estimate of the enemy, availability of resupply, replacements and maintenance, and casualty treatment are important force performance parameters, which are provided by the support forces.

Slide 1-1 — REQUIRED THEATER MODEL CHARACTERISTICS

Combat Force Representation
- Command Organization
- Reflect Battalion Mix
- Reflect Weapons Capabilities
- Include Combat Support Forces

Non-Combat Force Representation
- Communications
- Intelligence
- Logistics
- Medical
- Transportation

Command and Control
- Command Unit Decisions
- Reflect Doctrine
- Reflect Constraints
- Reflect Prior Battle Experience

Combat
- Produce Results of Combat
- Record Expenditures®consumption

Threat
- Reflect Unit Organization
- Tactics and Doctrine

Nuclear Forces
- Effects of Nuclear Weapons
- Employment Tactics and Doctrine
- Reflect Constraints
- Interactions Conventional®nuclear

Perhaps the most important operation to be represented in the theater model, and the most difficult, is the command and control process. It is here that the capabilities of all force elements converge to influence the course of the war. The information items I just discussed go into making command unit decisions and carrying them out. Current doctrine of force employment must be represented. Any constraints due to shortages and political or military policies must be reflected. It is also important to consider previous experience, that is, results of recent combat, inertia of combat, morale influences, and breakpoint considerations.

Combat forces, noncombat forces, and command and control combine to generate combat situations at some level of resolution in the model. Now, I'd like to digress here because General Welch talked about hierarchical representation in models. This is a place where there is a possibility of the hierarchical representation in the theater model. The theater model can generate combat situations. If you can fit the combat situations into high resolution models or games, this is the place to do it.

These combat situations are defined by the particular combat units and weapons present on each side, the missions of the forces opposing each other, the amount of combat support, and, of course, the terrain and environment. Now, this applies to all units on both sides. These situations reflect the decision process of the commanders with the resources they have available, the knowledge they have, and the normal response to such a situation. The next step is to generate the results of combat for every situation generated, the losses, the movement of the FEBA, the consumption of supplies, and the expenditures of ammunition. These items, of course, may be replaced with whatever is available from stocks, regulated by the capability of the logistics forces.

Now, I put "threat" down merely to note that our opposing forces may not operate as mirrored to ours. They may have organizational differences, as well as tactics and doctrines that are totally dissimilar to ours. These must be represented.

As I said earlier, the nuclear planning process is different from the conventional force process. It emphasizes weapon systems effectiveness against expected target complexes, investigation of employment tactics and strategies, and conventional nuclear interactions. Through all of these, political and military constraints need to apply. We need to consider the effects that the presence of nuclear weapons has on conventional war, even though the weapons are not employed.

Now, the Army is using theater models to address these force planning areas:

- Combat force requirements — CEM, ATLAS, IDAGAM
- Non-combat force requirements — CEM
- Readiness of current forces — CEM
- Sustainability of combat forces — CEM
- Nuclear forces — NUREX, COMBAT II

In combat force requirements, the CEM has been used for planning of weapon system/force mix analyses in the long-range period. Studies that were performed in the past, such as the CONAF studies and the conceptual force design studies that the Army did a few years ago, were the beginning of the use of CEM for this purpose. ATLAS is a much higher aggregated model. It is used annually to identify required numbers of divisions for objective force planning. This is a rather gross estimate, just trying to come up with a total number

of divisions, say six years from now, that would be required, to stop the enemy threat, the enemy penetration. ATLAS is not a good model for combat arms mix, but it is fast and as such can be operated in a requirements mode. That is the reason it's still in being. More recently, IDAGAM has been employed in a few studies of air support requirements.

An annual Army study called Total Army Analysis is conducted to determine the noncombat force requirements. The CEM is used to support this study. Actual deployment plans are used to phase in units to Central Europe. The CEM warfighting simulations are then conducted to generate data for replacements, consumption, and data on the pace of combat activity. These data are then combined in a further model with workload and allocation factors to generate support force requirements. Here we're also talking about five to six years from now.

Another Army study called Omnibus is also conducted annually and examines the readiness of current forces. The CEM is used here to measure the combat capability of the current force in its actual condition, and then to measure effects of correcting shortfalls in items such as replacement personnel and equipment, supplies, maintenance capabilities, prepositioned material, and so forth. Deployment or mobilization delays are also reflected by the arrival times of units in the combat areas.

The CEM also finds application for studies of sustainability, principally to generate wartime replacement factors for items such as ammunition, tanks, APCs, helicopters, and so forth. These studies are also run far out in the future, and these particular simulations are really sometimes less realistic than those that have been conducted for Total Army Analysis and for Omnibus. This is principally because replacements and resupplies must be unconstrained in this kind of operation in order to determine what is really needed.

Very little is done with theater models in nuclear planning. CAA, the Army Concepts Analysis Agency, employs NUREX/NUREM models in more or less single shot looks at weapon requirements. One model generates target complexes and the other matches the weapons to it. Combat II, a model developed by BDM, was recently employed by BDM in support of an Army study of theater weapon requirements. The Army itself has no theater model to address the range of nuclear planning problems I described earlier.

Now, summing up our applications, you'll see that we use the CEM more than any other model. This is because its characteristics match the required characteristics better than any other model. I was thinking awhile ago of this when General Welch was talking about paths. It just so happens that the path that you need to follow is the path that matches the requirements for the planning process. Any model that is not developed or not matched to that process is really going to find little use in the Army force planning process.

In the process of developing models like this we run into a real problem of trying to measure effectiveness of alternatives. This is a problem we haven't addressed adequately, but it's one that we should face. As you put decision processes in the model, the model sometimes leads to rather baffling results. That is, you would think by improving some capability you would see an increase in the force effectiveness, but sometimes you don't. We discovered this a number of years ago, we called it noise then — I think somebody has termed it the nonmonotonic effect. This is something that has to be recognized: if you increase the capability on one side, the other side can react to this, and he may take different actions than he took before, and therefore sort of countermand some effectiveness. It means that you have to look at a range of results. You can't pick one point and look at the result and say this is the way it should go. You have to look over the range of uncertainties. And this increases the analysis problem by at least the quantity of runs that must be made.

There is also the tendency here to measure end results of a battle. You start the war and then after so many days you look at the result — where the FEBA is, where the losses are. These sometimes are affected too heavily by the start point of the war and how the forces were deployed. We need some better way of measuring this, some way of averaging the force performance spaced over the spectrum of situations that have been generated, to assess how this particular force performs better than another force. Or, when we're trying to address shortfalls in our capability, as we improve these shortfalls, we must see what this effect is. And, I'll leave you with those problems.

Dr. Kapper: The next speaker will be Mr. John Shewmaker. He'll be talking about the Navy, and, I think also about the Congressional Budget Office (CBO).

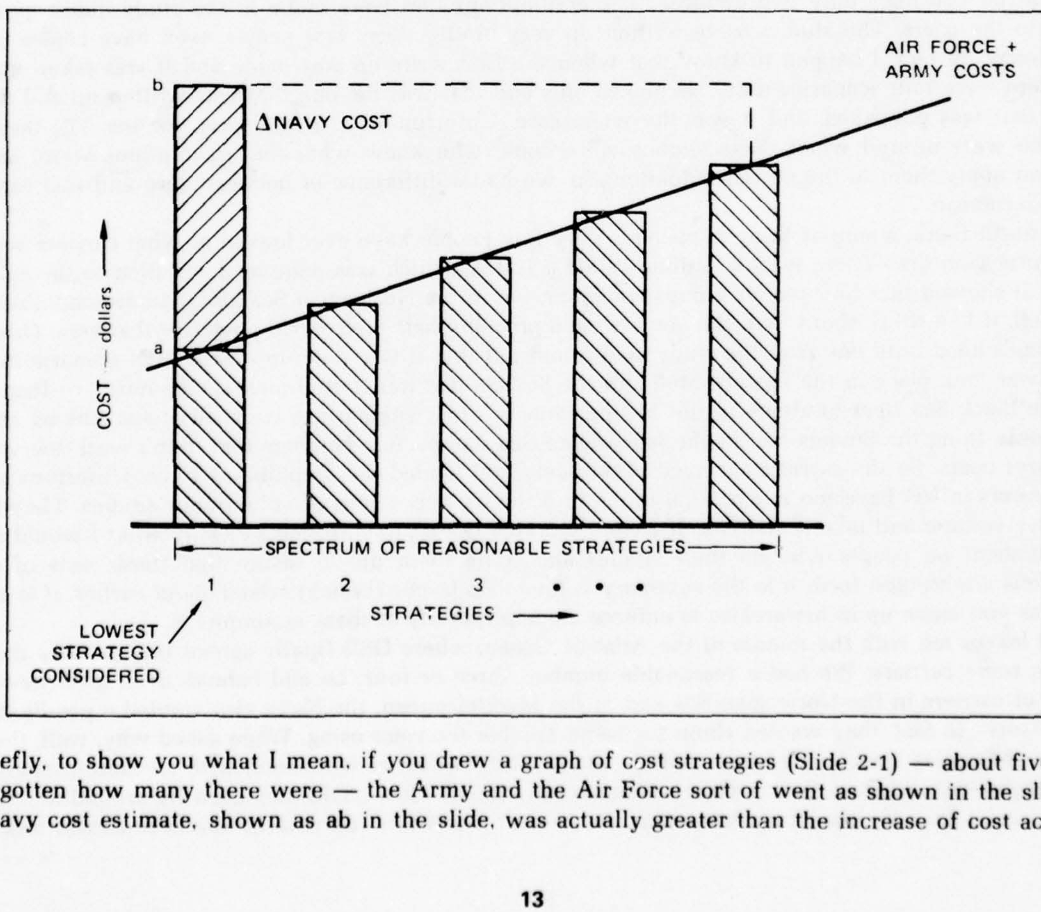
2 — Theater-Level Gaming in the Formulation of Plans and Policies (Navy)

MR. JOHN SHEWMAKER
Congressional Budget Office

PRM-10 and naval force requirements. A case of "lost" assumptions.

Mr. Shewmaker: Perhaps I'd better give a few reasons why somebody from the Congressional Budget Office will be talking about the use of theater level games in Navy force planning. I did just retire from the Navy this summer and went over to CBO, which helps explain a little bit of it. Somewhere along the line I was somewhat involved in developing some Navy models and studies — WAR AT SEA (NOW) and MEPEQ — that some of you may have heard of. And I was recently in OSD doing a little bit of Navy force level planning. So when Larry Low called me, and as I've listened to the words so far today, it seems to me to be appropriate that if we're going to speak of the use of gaming and force level planning for the Navy, a prime topic perhaps would be PRM-10. For those of you who keep track of those sorts of things, this was a study conducted for the National Security Council this spring in which Navy forces in particular were the cause of some modest controversy. In fact, the difference in cost between the forces derived by OSD compared to those derived by Navy for the lowest strategy considered, were greater than the cost for the Army and the Air Force forces across the spectrum of reasonable strategies.

SLIDE 2-1 PRM-10 COST REPRESENTATION



Briefly, to show you what I mean, if you drew a graph of cost strategies (Slide 2-1) — about five or six, I've forgotten how many there were — the Army and the Air Force sort of went as shown in the slide. The delta Navy cost estimate, shown as ab in the slide, was actually greater than the increase of cost across the

strategies. Yet, both of us used theater-level gaming to a great extent in deriving our forces. In fact, none of us argued much about what the results of the studies meant, how they should be used, and what they were telling us. So, I thought it might be appropriate then to give you just a few words on how all of this came to pass.

Well, PRM-10 started out with a concepts group which delivered to the force planners in Op-605 for the Navy and JCS and to myself in OSD, a group of strategies. We had strategies for NATO wars. We had strategies for wars in consonance with the NATO conflict and outside of the central front. We had intervention, we had East Asia, we had all sorts of different conceptual U.S. strategies in numerous alternatives. The planners, of course, tried to look through all of these problems and come to some reasonable estimate of force requirements. We had a lot of studies to back us up — the MEPEQ series that looked at the Mediterranean, which was of course, of great interest to the Navy, followed by NARAC-G, the current capabilities study, SEA MIX I, which has been practically a bible for Navy force planners, the CAPSTONE/CAPLOC series and the JCS, all spoke to what NAVY forces might be able to do in the context of a case one scenario.

Well, when it was broken down into theaters, the JCS and OSD pretty much agreed, about the intervention forces, perhaps because there hadn't been too much study done and therefore military judgment was primarily used and we agreed. In the Pacific and Indian Ocean theaters, similar agreements took place. With not too much action going on, a few carriers and some additional forces should be able to handle it. It was when we came to the Atlantic theater, which analysis concentrates on, that the parting of the ways took place. The Mediterranean, of course, is perhaps the best example. OSD asked the first question, I guess, as to why we wanted such a lot of naval force in the Mediterranean during a NATO war. Everything was going on in the central front. We had F-15's and we had Army and why should the Navy be terribly involved in the Mediterranean? In fact, some of the strategies indicated that perhaps our allies could defend there. The Navy response was very clear — the allies really aren't that capable. The war could have started there and we should be able therefore to put something in there to support them. How much? Well, it turned out in the MEPEQ series, which was carried through, there were some very important assumptions made. It turned out that if you need carrier task forces for the Navy to win that war, if the Soviets have access to the land bases on the Littoral of Africa, and the carrier groups are boxed in between, say Cyprus, Syria, and Egypt, within range of tactical air, it would take at least four or five carriers to try to win the war. If you delete any of those assumptions, for example, if the carriers start from Gibraltar and fight their way in, you don't need that many, maybe you need only two or three. The assumptions that were made in the study didn't quite filter through to the users. The studies were written up very briefly. Very few people even have copies of those studies today. In fact, I happen to know that when the first write up was made and it was taken up to the CNO, there were four scenarios used. He picked one, and that was the one that was written up and that was the one that was published, and it was the worst case. Unfortunately, in this case, for the 605, there were some who were around when those studies were done, who knew what the assumptions were, and who could then apply them to the current situation. So, we had a difference of between zero and five carriers in the Mediterranean.

For north flank, a similar thing happened. Very few people have ever looked at what carriers would do in the Norwegian Sea. There was something called a J-MIX, which was done as a variation to the SEA MIX scenario. It showed that four carrier groups could survive in the Norwegian Sea and land an amphibious task force. Well, if you think about that, the Soviets have probably half their naval power in that area. One might scratch one's head until one read the study, and found out that it was done in a SEA MIX scenario in which the first war took place in the Eastern Med and the Soviets had transferred most of the northern fleet bombers to the Black Sea fleet to attack in the Mediterranean. Their submarines were all at sea sinking convoys, and the only thing the Soviets had in the Norwegian Sea were a few bombers that didn't work too well and some patrol boats. So the carriers survived very nicely and landed the amphibious force. Unfortunately, the force planners in 605 have too much to do to dig into the backup details of voluminous studies. They look at a summary volume and take it and run. If I were working there I'm sure that's exactly what I would do. But it is incumbent on people who do their studies and write them up, to insure that these sorts of critical assumptions are brought forth into the summary volume. As Jasper (Welch) talked about earlier, it is a tough problem as you move up in hierarchies to enforce the applicability of these assumptions.

That leaves me with the middle of the Atlantic Ocean, where OSD finally agreed that perhaps the Navy could use some carriers. We had a reasonable number, three or four. Lo and behold, even after having put numbers of carriers in the Norwegian Sea and in the Mediterranean, the Navy also wanted a goodly number in the Atlantic. In fact, they wanted about the same number we were using. When asked why, well, they said they were defending the convoy lanes from air attack. But we sort of wondered, with all these carriers up in the Norwegian Sea, why are the bombers coming around them? How could they even fly around them? Well, we never really got very good answers to that. As a result, we wound up with estimates of carrier task forces

ranging from six to sixteen carriers — in some cases from nine to twenty-four. By the time you bring in the carrier costs, for the Navy, that dominates the whole picture.

605 had very good reasons for structuring their forces the way they did. In fact, given their assumptions, given the way they looked at the studies, they had a very logical, straightforward argument. They had a very easy time convincing Secretary Claytor that this was the right approach. What happened? Of course, PRM-10 went over to NSC and Secretary Claytor was in dismay about the whole thing, as was Russ Murray, and they agreed to do a study to find out what it is the Navy should be capable of doing in a NATO war. The models were not in question. The models have been run and we know the output. People don't argue about that very much. The question is, what's the scenario, what are the assumptions, what's the Soviet strategy, what mission should the Navy be capable of accomplishing? Once agreement on these is reached the theater level of games can be used to much better effect.

Let me leave you with just a couple of thoughts on what I see as the direction theater level games should go, for the Navy at least. In the PDM language this year, the Secretary of Defense has asked the Navy to look at deriving a Navy from the bottom up; peacetime deployments, what are required, how many forces do you need; intervention, what do you need to intervene in reasonable scenarios, and to define reasonable scenario. Then, look at how many more units of equipment would you need to fight a NATO war. There are those of us who think that the Navy is bought for case four and five, or at least the expensive portion of it, and the Secretary was interested at least in getting this sort of progression of analysis to see what it would look like. Would this justify the Navy forces in a better fashion? Well, how many theater level games, or how many models for gaming, treat contingencies in the Navy theater? When contingencies are considered, this begins to involve other services than the Navy because soon we have Air Force wings, strategic mobility, Army divisions, Marine Corps, and amphibious ships and carriers. What are the requirements in the contingency region? I don't think that the games are set up to really deal with those requirements today.

The second important point I'd like to emphasize relates to what strategy the Soviets play in these games. Almost every study that I've ever had anything to do with has optimized Soviet strategy against the measure of effectiveness used in the study with no relation whatsoever to anything the Soviets ever said about how they're going to use their naval forces. Now, that just is a major disconnect to my way of thinking.

A couple of JCS studies have tried to use a reasonable Soviet strategy but even then they only used one measure of effectiveness, ships sailed. My question is, what happens if we employ our forces to defeat the Soviet strategy? For those of you who may be interested, their first goal is to protect their SSBNs, for very good reasons, since we're particularly good in submarine-launched ballistic missiles. So, the problems that I see then with the Navy models are, (1) getting the assumptions *before the* decision maker, (2) looking at some contingency models, and (3) employing the Soviet naval forces sort of the way Gorshkov thinks they should be used.

Dr. Kapper: Although Gene Porter is down here to speak as the fourth speaker, he's being ably replaced by Bob Schneider. Bob, looking forward to hearing your presentation.

3 — Theater-Level Gaming in the Formulation of Plans and Policies (Planning, Analyses and Evaluation)

MR. ROBERT SCHNEIDER
OSD-PA&E

A spectrum of techniques used in PA&E for force balance assessments. The issues of advocacy and lagging data base development.

Mr. Schneider: I'll try to keep my remarks fairly short. My old boss, Gene Porter, was to be here, and I think the situation that brought the fact that he isn't here is indicative of how we in OSD P&E view models, the use of models, because we have, I think, a rather peculiar problem. It's very much illustrated by the Indian who was drafted into the Army during World War II. When he was able to get out of basic training and back home on leave his friends asked him what he thought of the Army, how he liked it, and he said, "there's too much tootin' and salutin', and not enough shootin'." I think the Planning Analysis and Evaluation Group, within OSD, view ourselves very much as being up on the line in

combat and very much concerned about the shootin', and because of that we don't have very much time to get back into the salutin' and the tootin'. Since we are very much oriented to responding to whatever the need is at the moment, we will pick up any rock that is close by, or missile if necessary, in order to destroy that man or that tank, or whatever it is, on the hill. Consequently, in OSD theater level models are used primarily for force balance comparisons, and I would say that many different techniques are used to compare the capabilities of the opposing forces. The primary objective is to help defense decision makers in allocating resources.

The basic difficulty faced by both analysts and decision makers is that no single technique or tool is always adequate, so we have that problem. We have another problem, and that is the problem of "proponency." Typically, when everyone is fighting for the dollar, there is an adversary relationship that develops, and as a result one proponent is rather suspicious of another proponent's model, especially if it's complicated and they can't see what's inside of it, because assumptions can be hidden, just as was already brought out by John Shewmaker. Very honest appraisals or opinions can be different but the assumptions on which they are based can be hidden.

Now, the assessments of opposing forces, therefore, must be developed from a combination of different force capability indicators. Each of these indicators will have a certain advantage or disadvantage and we must take that into account. So, in practice, this means that defense decision makers receive a wide spectrum of force balance assessments ranging from very simple numbers counts of men and equipment all the way to outputs obtained from very sophisticated war gaming models. And, of course, this is all sprinkled with mature military and civilian judgments. So, really, we get involved in three different levels of analyses. The first one is one in which the numbers of tanks, aircraft, men, and so forth, are derived, because they're easy to derive and they're also easy to present. In many cases, they're also the easiest to understand, and I might say they're also the most difficult in which to hide assumptions. However, number comparisons may be very misleading if the counting rules are not clear and explicit. In addition, they imply that the opposing forces should be mirror images of each other, when, in fact, forces may be structured and equipped to satisfy very differing demands, as John also brought out in looking at Soviet strategy. So, number comparisons neglect force interrelationships; they neglect the quality of different weapons, which are important.

Now, at the other end of the spectrum, war games generally attempt to model force interrelationships and to predict some sort of an outcome of the hypothetical conflict. They're especially useful in assessing the relative importance of different factors/weapons and in understanding how these factors/weapons influence the outcome. However, due to the number of implicit and explicit assumptions in most models, predictions of the conflict outcome are often discounted by decision makers. One of the big problems in modeling, and from the point of view of people who are involved in doing large aggregations, is that model development is way ahead of the input data availability. So, for example, if I model a war in Europe, and only take into account half the tanks that are involved, if a decision maker knows I'm only accounting for half the tanks, he has difficulty understanding what relationship the results have to his problem.

The third technique that we use is a compromise between these two. This technique is known as the WEI/WUV approach to calculating armored division equivalents. Now, this has received increased attention in OSD but not necessarily other places, and it's increasingly being used as part of the theater-wide ground force assessments. It is more complex than just a number counting exercise, but it is a lot simpler than using the theater model or the war game. Although the technique suffers from many of the deficiencies associated with the simple numerical comparisons and the war games, decision makers have found the technique to be a useful indicator in ground force balance assessments, and it is being more widely used.

Of course, all three of these techniques are still deficient because they neglect a lot of the intangibles like leadership, training, morale, and in many ways, surprise, weather, and so forth. But, of course, decision makers know that. So, in OSD we'll use whatever is at hand and whatever is needed in order to get the point across at the particular moment.

Let me just review a few of the uses of the models during this past year. PRM-10, as many of you have read in the newspapers, had many, many different strategies. It was done in a hurry, and, consequently, those in PA&E, and I'd say in OSD in general, didn't have time to go back to anything with which to run an analysis or a problem. So they took whatever was at hand. In these gross strategies, where decisions are made between very large differing strategies, people tend to say, "Well, when the Army did this, and the Air Force did this, these models generally agreed in this area, therefore, we'll pick that answer." I can't go into much detail here, but there wasn't much question about what the models came out with, or even the assumptions. The conversation was at even a higher degree of aggregation than the problems that those models were being used for, so the models were just taken for whatever purpose they might be used. Of course, you and I know

that when the chips are down models often are used just for political reasons. If it comes out with a right answer, no matter how it got the answer, that's the right answer and so it is used.

Another use of models was in this past issue cycle in which on the logistics side of things there was disagreement about ammo rates. Every year the Army does a large study using theater level models to come up with ammo rates. Well, because that has become an issue in this year's program, the analyst in OSD said, "We'll go ahead and use that model, but we'll change the assumptions." So, in fact, all that happened was that the analyst asked the Army to run the model just as they would normally run it except to change certain assumptions and see how it came out. Then the analyst used the difference between the two answers, based on those assumptions. There, the model was accepted and it was used in this particular area as a sensitivity analysis. For determining how many tanks we should have we conducted a war reserve study on tanks. For that study, P&E used a Lulejian model. The analyst completely controlled it, ran it from beginning to end, and used it in a sensitivity analysis to decide how many tanks were needed in war reserves. Since accuracy wasn't really as important as getting a general feel for what the difference was going to be between stopping at a particular FEBA or being able to move out on the offensive, that difference was great enough so that it gave the kinds of answers that the analyst needed.

A third approach to the use of models within OSD has been illustrated in a recent study by the Military Committee Special Study Group (MCSSG). Because this was an international study in which the Germans, the British, and the Americans each were using different theater level models to address the same problem, OSD's function almost became political in a sense because we wanted to make sure that the assumptions that were used would not drive the models to the wrong conclusions. So our concern wasn't so much in the running of the models as it was in the assumptions that were used that could then be used to make the result come out contrary to U.S. interests. That may be a revelation to some of the people that were involved in the study.

So, we get involved in modeling in many different ways, but I hope that this illustrates to you that theater level models are the only way to go in order to do certain things. Very often our big problem is failing to attach the model and its use to the problem at hand, we need to concern ourselves with making proper use of the models that we have and connecting them to the problem at hand.

Dr. Kapper: The next speaker will be Mr. Andrew Marshall, Office of the Secretary of Defense and Director of Net Assessment.

4 — *Theater-Level Gaming in the Formulation of Plans and Policies (Net Assessment)*

MR. ANDREW MARSHALL
OSD — Net Assessment

The lack of theater model utility in Net Assessment.

Mr. Marshall: Having heard the previous speakers I find myself perhaps in the position of delivering the most negative or most pessimistic view of the current state of modeling, at least as seen from the use to which I've been able to put modeling. First, however, let me say I don't see myself as being in the force planning business. I'd say I'm in the diagnostic business, leaving it to others to decide what might be done or ought to be done. I'm involved in trying to do analyses of various military balances with the objective of calling to the Secretary of Defense's attention emerging problems, adverse trends, or opportunities that somebody ought to do something about.

When I first got into this business in 1972, one of the first things I did was to call in some people and try to survey the models that existed to see whether there really was anything that was going to be of much use to me. The answer was, there wasn't. Now, I think it may be appropriate for me sometime soon to do another survey and see whether things have improved sufficiently that there would be a different answer. But I would say that in none of the work that I have done have we run a single model or asked anyone for very elementary calculations. Indeed, in the papers that we write we often give the results that other people are giving on trends of various indices, often sometimes to give a description of what has been happening on primarily the U.S. and Soviet sides. Often this is done to say to the Secretary of Defense, "This is what a lot of people are going to be looking at, this is what is going to be presented to the public, but here are six reasons why those are poor ways of looking at this balance, and here is what you need to know about things that are going on or uncertainties that nobody else is going to call to your attention." As I say, to a large extent we have focused on just describing the current state of affairs, the past, how we got to it, the major asymmetries in the forces on the two sides, both in doctrine or the force posture, and regarding many other factors.

One of the things that concerns me most about the available models, although I haven't looked recently so I might conceivably be incorrect, is that they don't allow some of the major asymmetries that exist in tactical doctrine to be introduced. I have the sense that much of the modeling has been done for the force planner with the kind of simplifications that perhaps can be introduced for his purposes, but I'm also particularly concerned by the lack of treatment of command, control, communications (C^3) in its broadest aspects in models. I don't mean just the communications nets, I mean the whole structure of the decision making on the two sides, which might be different, or the vulnerabilities of command and control, or its susceptibility to deception or disruption. One of the reasons that it particularly concerns me, one especially from doing some comparative analysis for U.S. and Soviet military doctrine is that the Soviets give high priority to attacks on C^3. In contrast to our doctrine, which is heavily or almost exclusively attrition oriented, the Soviets focus on ways either through deception, surprise tactics, physical attack, electronic warfare, to disrupt the central nervous system of the other side rather than exclusively looking at out-muscling them. So, if I have that view of what the combat might be about, or a major element of it, models that do not allow you to treat that doctrine don't offer a lot of confidence.

And I must say in this connection that, looking back over military history it concerns me that if you take most of the indicators that people use, even WEI/WUV's and all of that, and apply them to something like May of 1940, looking at all those numbers, the French, British, Dutch, and Belgians ought to have at least held the Germans. Instead they were devastated in a few days to a few weeks. What the Germans had was not better equipment or more of it. They had a better idea of how to use it, better leadership, and a better overall plan.

Again, it worries me that in the modeling that I see of the famous central front the war is assumed to start with a kind of routine advance by the Soviet forces into the NATO position. Now, again look at military history. In any major jump-off against an enemy, as far as I can tell, the attacker has focused on deception, ways of surprising the other side's special operations, to the rear or to hit specific key points. All of which have had enormous leverage in historical cases. The models are just not up to treating that kind of thing, at least at present. Instead, they treat the steady state case in which engagements are sort of going as planned by the two sides. I think for all of these reasons I have felt that the current modeling or indices that people have are not really descriptive of almost any military balance that I really know of, including the strategic balance, which one would set aside for a separate discussion.

On the other hand, I do not mean to be overly discouraging. I would hope that, say, in ten years or so there really would be models that would be much more useful. I don't know whether that's going to be the case, but I would encourage people to work toward that goal.

Dr. Kapper: I saved this last spot for myself, because I do have a couple of points I would like to touch on.

5 — *Theater-Level Gaming in the Formulation of Plans and Policies (JCS)*

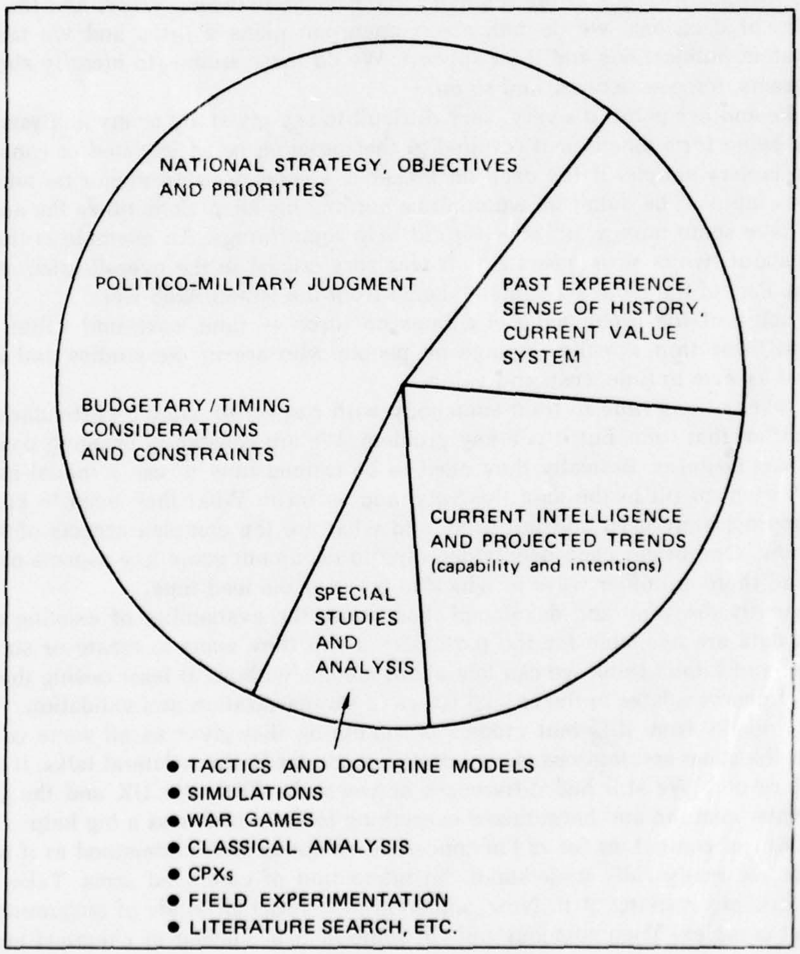
DR. FRANK KAPPER
JCS/SAGA

The impact of analysis on the decision process, how models are used by the JCS — and some related problems.

In this session we were to look at the use of theater combat models in the decision making process and at the weaknesses, if you will, of those particular models.

Models have to be kept in context, and what they produce has to be kept in context as well (Slide 5-1). The decision maker uses computer results as input to his decision making process, not as his basis for a decision. He also uses quite a few other things. He uses the capability and intent that he gets from the intelligence side of the house, and he uses his own knowledge of military history. I think if you look at what models produce, unless you keep this in perspective, you're going to lose track of a model's relativity.

SLIDE 5-1 KEY FACTORS INFLUENCING THE POLICY/DECISION MAKER
(A Subjective Viewpoint)



Again, with respect to theater combat models — and they're very important — and it's like the ad on TV, "if we don't have them we're out of business." Well, it's not quite that severe, but models are the sine qua non to do a lot of what we do because we can't run wars very easily or cheaply, or quite as fast.

Models are a means to an end; models are tools. They are not the end objective in themselves, except to the model maker and the model designer at a particular stage in the model's development. From a users standpoint, though, which is what this session is about, they are tools, and very helpful ones. But I think that models have got to be put in the more meaningful context of studies and analyses. We don't just run a model or models — we do it to study something or to analyze something, some kind of phenomena. So, again, it is important to keep in mind that models are an input to studies and analyses.

Studies and analysis have many purposes, but their main purpose is to understand the phenomena — that's really the most important thing. If you can predict, you can control, and do all those other wonderful things — superb — but the primary purpose of models is to gain an understanding of very complex phenomena.

Now, in talking about where the JCS uses models relative to types of decisions, in a theater context, we do use them relative to the support of the Joint Strategic Objectives Plan (JSOP), for the Joint Force Memorandum (JFM) and in special analyses. One of the reasons *why* we use them, of course, is to determine the relative risk for particular force postures. We don't determine risk in a vacuum; it's relative to a threat. So I think that's one of the most important things vis-à-vis model use in the JCS.

In the JCS we also do a great many special analyses in which we use theater combat models. The mutual balance force reduction (MBFR) series of analyses use theater combat models. The Short Pact Attack Study, or scenarios, investigated the ability, and the desirability if you will, of the pact to initiate attacks given certain short duration preparation times. We use models a great deal in the CAP series of studies for logistics, and mobility, CAPLOC, CAPSTONE, and so forth. In addition, we used models in the Military Committee Special Study Group (MCSSG) to look at the relative force balance between NATO and the Warsaw Pact. So, in terms of the types of decisions, we do talk about operation plans a little, and we talk about the joint strategic planning system publications and their support. We do these studies to identify shortfalls and weaknesses in plans, programs, force structures, and so on.

Now, let me make another point. It's very, very difficult to say my study or my analysis really resulted in this particular policy being formulated, or it resulted in that program being initiated or cancelled. The reason why it's tough to do is very simple. If the decision maker is a good decision maker he takes that study and analysis and uses it as input. If he didn't he wouldn't be earning his keep. Sometimes the analyst can identify where his work did save some money, or where it did help some things. An example is the shelter program study which we did about two or three years ago. It was very crucial in the overall orientation for NATO to go to aircraft shelters. Part of the data, incidentally, came from the Israeli/Arab war.

Now, what are some of the problems? Let's focus on three — time, cost, and value. That's kind of a general overhead, but these things really impinge on people who are in the studies and analysis business. And, there are several aspects to time, cost, and value.

For example, it takes a long time to train somebody with respect to using a particular model. There are ways in which to shorten that time, but it is a key problem. We are constantly bringing people in and out of the studies and analysis business. Basically they need to be trained how to use a model more so *after* they learn to punch this button, to fill in the data this way, and so forth. What they need to know are its weaknesses, where is it appropriately used and not used, and what are the complex aspects of the model that as users they should know. One bright chap used video-tape to document some key aspects of his algorithm — that's a great idea. And there are other ways in which to get over the lead time.

Models are frequently designed and developed that ignore the availability of existing data bases, or ignore whether or not data are available for the particular action they want to create or simulate. There are data base weaknesses, and I don't think we can talk about models without at least raising the data base weakness issue. And this of course relates to the crucial issues of standardization and validation.

Comparability of results from different models is something that gives us all some concern. Assuming that we all start with the same assumptions, as we pretty much did in the trilateral talks, it is still possible to have vastly different results. We still had differences between the FRG, the UK and the US, and then the SHAPE Technical Center came in and harmonized everything for us which was a big help.

The phenomenology of combat, as far as I'm concerned, is not as well understood as it needs to be. In all honesty, I don't think we really fully understand the interaction of combined arms. Take a typical ground force situation, with one guy looking at it. Now, someone introduces all kinds of air/ground interaction and then it gets somewhat complex. Then someone puts in some tactical nuclear or chemical munitions and that

5 — Theater-Level Gaming in the Formulation of Plans and Policies (JCS)

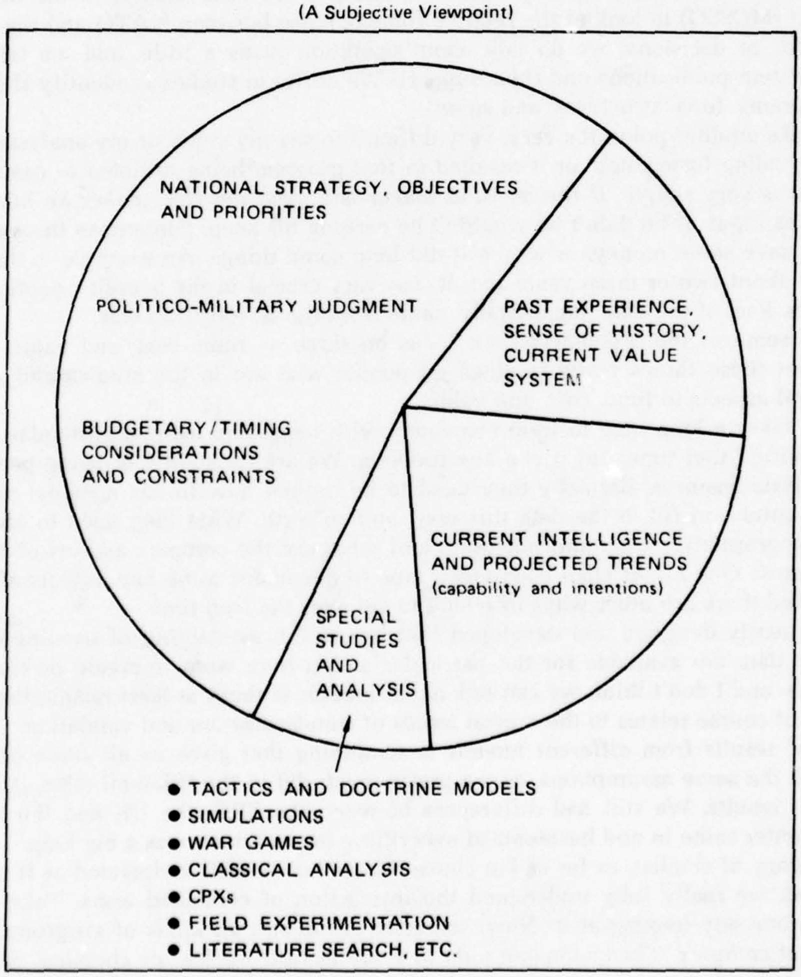
DR. FRANK KAPPER
JCS/SAGA

The impact of analysis on the decision process, how models are used by the JCS — and some related problems.

In this session we were to look at the use of theater combat models in the decision making process and at the weaknesses, if you will, of those particular models.

Models have to be kept in context, and what they produce has to be kept in context as well (Slide 5-1). The decision maker uses computer results as input to his decision making process, not as his basis for a decision. He also uses quite a few other things. He uses the capability and intent that he gets from the intelligence side of the house, and he uses his own knowledge of military history. I think if you look at what models produce, unless you keep this in perspective, you're going to lose track of a model's relativity.

SLIDE 5-1 KEY FACTORS INFLUENCING THE POLICY/DECISION MAKER
(A Subjective Viewpoint)



Again, with respect to theater combat models — and they're very important — and it's like the ad on TV, "if we don't have them we're out of business." Well, it's not quite that severe, but models are the sine qua non to do a lot of what we do because we can't run wars very easily or cheaply, or quite as fast.

Models are a means to an end; models are tools. They are not the end objective in themselves, except to the model maker and the model designer at a particular stage in the model's development. From a users standpoint, though, which is what this session is about, they are tools, and very helpful ones. But I think that models have got to be put in the more meaningful context of studies and analyses. We don't just run a model or models — we do it to study something or to analyze something, some kind of phenomena. So, again, it is important to keep in mind that models are an input to studies and analyses.

Studies and analysis have many purposes, but their main purpose is to understand the phenomena — that's really the most important thing. If you can predict, you can control, and do all those other wonderful things — superb — but the primary purpose of models is to gain an understanding of very complex phenomena.

Now, in talking about where the JCS uses models relative to types of decisions, in a theater context, we do use them relative to the support of the Joint Strategic Objectives Plan (JSOP), for the Joint Force Memorandum (JFM) and in special analyses. One of the reasons why we use them, of course, is to determine the relative risk for particular force postures. We don't determine risk in a vacuum; it's relative to a threat. So I think that's one of the most important things vis-à-vis model use in the JCS.

In the JCS we also do a great many special analyses in which we use theater combat models. The mutual balance force reduction (MBFR) series of analyses use theater combat models. The Short Pact Attack Study, or scenarios, investigated the ability, and the desirability if you will, of the pact to initiate attacks given certain short duration preparation times. We use models a great deal in the CAP series of studies for logistics, and mobility, CAPLOC, CAPSTONE, and so forth. In addition, we used models in the Military Committee Special Study Group (MCSSG) to look at the relative force balance between NATO and the Warsaw Pact. So, in terms of the types of decisions, we do talk about operation plans a little, and we talk about the joint strategic planning system publications and their support. We do these studies to identify shortfalls and weaknesses in plans, programs, force structures, and so on.

Now, let me make another point. It's very, very difficult to say my study or my analysis really resulted in this particular policy being formulated, or it resulted in that program being initiated or cancelled. The reason why it's tough to do is very simple. If the decision maker is a good decision maker he takes that study and analysis and uses it as input. If he didn't he wouldn't be earning his keep. Sometimes the analyst can identify where his work did save some money, or where it did help some things. An example is the shelter program study which we did about two or three years ago. It was very crucial in the overall orientation for NATO to go to aircraft shelters. Part of the data, incidentally, came from the Israeli/Arab war.

Now, what are some of the problems? Let's focus on three — time, cost, and value. That's kind of a general overhead, but these things really impinge on people who are in the studies and analysis business. And, there are several aspects to time, cost, and value.

For example, it takes a long time to train somebody with respect to using a particular model. There are ways in which to shorten that time, but it is a key problem. We are constantly bringing people in and out of the studies and analysis business. Basically they need to be trained how to use a model more so *after* they learn to punch this button, to fill in the data this way, and so forth. What they need to know are its weaknesses, where is it appropriately used and not used, and what are the complex aspects of the model that as users they should know. One bright chap used video-tape to document some key aspects of his algorithm — that's a great idea. And there are other ways in which to get over the lead time.

Models are frequently designed and developed that ignore the availability of existing data bases, or ignore whether or not data are available for the particular action they want to create or simulate. There are data base weaknesses, and I don't think we can talk about models without at least raising the data base weakness issue. And this of course relates to the crucial issues of standardization and validation.

Comparability of results from different models is something that gives us all some concern. Assuming that we all start with the same assumptions, as we pretty much did in the trilateral talks, it is still possible to have vastly different results. We still had differences between the FRG, the UK and the US, and then the SHAPE Technical Center came in and harmonized everything for us which was a big help.

The phenomenology of combat, as far as I'm concerned, is not as well understood as it needs to be. In all honesty, I don't think we really fully understand the interaction of combined arms. Take a typical ground force situation, with one guy looking at it. Now, someone introduces all kinds of air/ground interaction and then it gets somewhat complex. Then someone puts in some tactical nuclear or chemical munitions and that

really creates a lot of complexities that I really don't think we understand. I think we should try to look at the basic phenomena and try to get a better handle on the essentials. I don't think we do enough in this regard.

Also, I have a feeling that we often use model studies and analyses to make up for quantitative weakness, vis-à-vis NATO/Warsaw Pact. I think we can play qualitative games in this area only so long before there will be some severe prices to pay.

Lastly, there's a lack of frequency in the communication that we have with our allies, particularly our German and British allies. I think we really need to pull our allies into our modeling business, into our studies analysis, and gaming business, if you will. They have learned a lot of lessons and have perspectives which can be of great benefit to us all.

Dr. Kapper: Our next speaker will be Jim Starkey from CIA, who will talk about theater-level gaming from their standpoint.

6 — Theater-Level Gaming in the Formulation of Plans and Policies (CIA)

MR. JAMES STARKEY
CIA

Initiation of a theoretical intelligence program using theater models.

Mr. Starkey: We are just beginning to get into theater level modeling in the CIA and our main interest, of course, is in trying to evaluate the effectiveness of military forces of foreign nations.

A lot has been done at CIA with regard to the effectiveness of strategic forces over the past few years and with good results, at least so the people believe who've been doing it. The sort of thing that we're interested in doing is to try to get a better estimate of the military effectiveness of forces in other areas of the world, particularly Soviet forces, and perhaps as a distant dream to try to develop on the Soviet side some estimates of their future force planning as those who have spoken earlier have described for our own side. One of the problems that our own force planners come up against all the time is that when they're planning for 1980 or 1985, the data they have about Soviet forces, for example, are not very good. Perhaps by getting a better understanding of the effectiveness of Soviet forces now, the sorts of things that the Soviets seem to feel are important to them, we can make some better estimates of how their forces may look, or how they may want their forces to look in the future.

A few models are either in use or under development at the Agency. The TACOS model is being used to study the penetration of air defenses by tactical air. The IDAHEX model is under development in a study of Soviet and Chinese forces in the Far East. We are using the UNICORN model, which is a weapons allocations model for nuclear, chemical, and conventional weapons, and we have under development the FIREARM model, which is being used to measure the ability of two sides of different size to project military forces to crisis areas in various parts of the world. As I said, we're beginning to get into modeling, and we're proceeding slowly and cautiously, as perhaps Mr. Marshall might advise.

Dr. Kapper: Our next speaker is Colonel Miles March from the Arms Control and Disarmament Agency.

7 — Theater-Level Gaming in the Formulation of Plans and Policies (ACDA)

Col. MILES MARCH, USA
ACDA

Theater-level models in arms control and arms limitation studies.

Colonel March: Some people might wonder what a dove is doing at this hawk conference. All right, here we go. I'm currently assigned to the United States Arms Control and Disarmament Agency (ACDA). Our interest in theater-level war gaming perhaps might best be answered by asking four quick rhetorical questions. As kind of a preface, I might say that I've only been with the Arms Control and Disarmament Agency for about nine months, so my knowledge of its functions is somewhat limited; also, of course, politics being what they are, the turnover in personnel has been pretty drastic down there, and the things that I've gotten together on ACDA's experience in this area are taken from the institutional memory represented by published things that I ran across, and the memories of the remaining people who are aboard.

The agency itself was formed in September 1961, and these are its basic charges:

- Conduct studies and provide advice relating to arms control and disarmament policy formulation.
- Prepare for and manage United States participation in international negotiations in the arms control and disarmament field.
- Disseminate and coordinate public information about arms control and disarmament.
- Prepare for, operate, or as needed, direct U.S. participation in international control systems which may result from United States arms control or disarmament activities.
- In addition to directing the agency activities, the Director will act as the principal advisor to the President, the National Security Council and the Secretary of State in arms control and disarmament matters.

Additionally, in recent times, we've been charged to do these things:

- Analyze annually every nuclear weapons program for which funding was being requested, other military weapons programs meeting designated monetary thresholds and certain other programs that might have significant arms control implications.*
- File reports indicating to what degree any significant arms control proposal made to or by the United States can be verified by existing national technical means and the continuing capacity of the United States to verify components of arms control agreements or treaties already in force.*
- Participate in the arms transfer review process in connection with the proposed sale or transfer of military technology or equipment to other nations.†
- The Director will also serve as the principal negotiator for the United States at the Strategic Arms Limitations Talks (SALT).‡

*Added by amendment to the original act creating ACDA.
†Added by amendment to the Arms Export Control Act and the Foreign Military Sales Act.
‡An additional duty of the incumbent director.

The statutory responsibilities and some of the tasks that ACDA might be asked to look at, using theater level war gaming, are as follows:

- Mutual and balanced force reduction option comparisons
- Arms control impact statements
- Arms transfer analyses
- Arms limitation agreements
- Risk analyses of verification assessments

In 1971, when the mutual and balanced force reduction (MBFR) options were being looked at by the Defense Department, its various services, and other agencies, ACDA also got in the act in a small way. It

brought on board ATLAS from GRC and had it modified somewhat because there weren't enough people to feed it; finally, they developed a FORPEM model, which is a synthesized ATLAS. It took out all because there weren't enough people to feed it; finally, they developed a FORPEM model, which as a synthesized ATLAS. It took out all the answers that ATLAS could possibly give you in terms of reality, but it allowed you to run a number of options through a common processing and get an order of preference rather than any close comparisons on which option was relatively better than another.

Slide 7-1 COMPUTER MODELS USED BY ACDA

Strategic Arms Limitation Talks (SALT)

- Strategic Exchange Models
  - AEM
  - RPM
  - SIR NEM
- System Capabilities
  - AIRPEN (SIR NEM)
  - ICARUS
  - PATH 87W/WA/WB
  - PROGRAM STRIKE
  - RANGE
  - AWACS SURVIVAL
- Strike Analysis
  - COMPARE
  - FOOTPLACE
  - GNPSUR
  - GNPUS
  - HARDSITE
  - PkPOP
  - FALLOUT (RPM)

Nuclear Test Ban Verification

- Cam
- Earthquake Hiding
- Network
- Seismic Event Process

MBFR and Arms Balance Issues

- ATLAS
- FORPEM
- NATO Deployment Model
- Static Force Ratio Model

Some of the models that we have on board at ACDA are shown in Slide 7-1. You'll notice as a small agency, with only about 240 people all together, the emphasis is on the hottest fire, and, of course, the strategic arms limitation talks have been that.

The agency has a lot of people who are closely involved with verifications. You'll notice that in the nuclear test ban verifications they go in for the idea that the Soviets *might be hiding* their nuclear tests in earthquakes, and they want to be able to detect, using national technical means, the size of nuclear bursts conducted in the Soviet Union. So we do a good deal of modeling in this area with testing stations throughout the world and the exchange of data. Again, in conventional theater modeling we're sort of limited to the ATLAS, FORPEM, a so-called NATO deployment model that was taken from a British model, and lastly, to the Static Force Ratio model, which I don't think is used at all any more, but we happen to have it (the latter was developed in OSD systems analysis and was another very quick processor for MBFR options).

I think that will give you some idea of what use ACDA might find for theater-level models. In the strategic field, we're a consumer as far as theater level gaming is concerned. We do investigations of things like Cruise missiles and the effect on varying degrees of hardness on silos — more or less the nuts and bolts rather than taking the whole thing and looking at it in a theater level context. We try to evaluate, with a small staff, the things that are put out by other people, and as such we hope to keep aware of the state of the art and monitor the developments that you people are making.

Panel Discussion — Session I

Colonel Grace: By way of introduction, I'm working in the Requirements and Programs Division at Headquarters, Marine Corps, as head of a little branch we're trying to get started which we call, strangely enough, Program Analysis and Evaluation. The purpose of our name is to identify what we're trying to do both to those within the Headquarters staff and to those outside the Headquarters. In addition, I'm one of the two Marine members on the follow-on study to PRM-10, irreverently referred to by some as the son of

PRM-10.

The Marine Corps, institutionally, is not very "at home" in the force planning business as it's conducted in the Washington area. We're more at home in the environment that we're talking about at this meeting. When we come into the Washington environment we're at a disadvantage for a couple of reasons. We're small — we represent something of the order of 10% of the total resources devoted to general purpose forces, if you throw in amphibious forces and a few other odds and ends. We're conservative — partly by nature and partly because of a lack of willingness to put a lot of resources into talent or computer time, if you will, or systems, that the larger services use to do these analyses that are to be an aide to decision makers.

We're also, in a sense, victimized by the process of force planning. I'm speaking positively now. I can see the other side. I don't want to throw rocks because every one of the problems I am going to mention I see as a challenge.

Phil Louer said that the models that don't fit the planning process will not be used. I think that is particularly appropriate to the Marine Corps. We're broken up in little pieces and our similarities are analyzed, if you will, in comparison with the larger services. We have people that fight on the ground and they look sort of like Army infantry divisions. We have high performance airplanes, which if they're flying off a land base, look sort of like Air Force high performance airplanes, and if they're flying off a carrier the only difference between those of the Navy is that they have USMC painted on the side instead of USN. And then we float around in Navy ships, only many times the Navy doesn't understand why we're floating around on those ships, especially when they see the price tag on those ships that could be directed to some other kinds of ships that they understand better. So, we're sort of working on the margin, if you will, of all the big guys, and one of the points that the little guy sort of takes as conventional wisdom is that when the elephants start fighting you get out of the grass.

On the other hand we like to think that we have some unique as well as substitutable functions to contribute to the general purpose forces of this country. The problem is how to explain them in terms that decision makers can understand. Most decision makers whether they only read it from history or experienced it themselves know about World War II, and they remember pictures of Marines charging up beaches. That's amphibious warfare in many people's minds. But, as John Shewmaker mentioned, there is a factor called contingencies — the "less than case one" scenarios in which a naval force, including embarked landing forces at some relatively small level of force, can make some important contributions if you can get them there, (strategic mobility) if you can project them ashore (depending on what the mission is, in what quantity and in what form you want to project them ashore) and if you can sustain them without the prior existence of some base structure and built in logistic support system. So, it's in situations like that that we feel, with the Navy, that we offer some valuable contributions.

Another factor that is peculiar to Marines — they are sort of offensive creatures, (in more ways than one). And I was particularly happy to hear Andy Marshall talk about the problems of asymmetries and the advantages the man with the initiative has. This is consistent with my military training — you're not going to win much if you remain back on your heels and let the other guy keep hitting you as you try to fend him off. So, at the operational level, we're sort of offensively oriented, but I recognize that at the policy and strategic level we have to be defensively oriented. So in the development of a hierarchy of models, how do you bridge that big gap? How do you, for instance, optimize a force like I described, a self-contained, however small, ground and tactical air combat force with sort of a built in strategic mobility and logistic support system? How do you optimize that for offensive action and evaluate it in that light yet fit it in within the generally defensive philosophical approach to all defense planning? As I see it, those are the kinds of problems that Marines face when they move into the Washington environment of both force planning and as a subset of that, modeling.

Mr. Hoeber: As a consultant, I'm not quite sure what I can say about users of models. But I really want to make one quite brief comment. I got the feeling this morning at least from some of the talks, that maybe the greatest danger is in believing our own models and our own numbers. I much appreciated our chairman's closing remarks, — that our particular tools feed into only a very small part of the decision makers equipment. This is a bad year to say it, but decision makers don't really operate on a zero base budgeting philosophy. They cannot start over and decide what the level of forces should be in force planning. I think the PRM-10 example was a great one. Depending on the assumptions, we could get anything from zero to sixteen as a requirement for the number of carrier task forces. Decision makers do operate marginally, and, if there is a contribution to be made in the present state of the modeling art, I would think it is in being able to do sensitivity tests, not all of them, because as Jasper pointed out, 2^{100} or 10^{30} is forever, and we can't do them all. But we can do a great deal in telling the probable effects of marginal changes, and that is a tremendously valuable contribution if we don't spoil it by suggesting that our numbers are the correct ones. We have trouble enough being sure that the assumptions that we make don't distort even the marginal sensitivities that we measure. But we can do a great deal, and I think many of the models I heard discussed this morning, some of which I'm familiar with, do give

very useful sensitivities to decisions that decision makers have to make.

Even if the absolute number we're talking about is a rate rather than a level, as in trying to establish consumption rates for ammo and attrition rates for equipment. The only good model of the last few years that I know of in this area is the model war of October 1973, and that suggested that all of the consumption rates that have been in models in use were much too low. I've seen some that have not changed since that war, but I would suggest that that's the only kind of model that you can take seriously and believe the results. Even then, you'd better be sure that the assumptions are consistent with the assumptions that make sense for the next war that you're planning since you can't run that kind of model very often. I guess what I'm saying is we ought to have a little humility and stick to what we can do and do best.

Dr. Linnenkamp: I'm presently working for the planning staff of the Federal Minister of Defense in Bonn, FRG. Before that I was in the Army staff of the research section for a number of years.

Basically, I would like to make one point as far as the presentations of this morning are concerned. I think that the approach of theater-level gaming tends to look at a very big system, war, or scenario development, for example, in Central Europe. This is certainly a very complex system which is to be analyzed by modeling in any form, simulations or manual war games or what-have-you. This overall view, which has to be taken in modeling theater level conflicts, produces at the same time, I think, a tendency to give relatively simple, or big, answers such as this war will last so many days or so many weeks, or this is a force compared to another force that is probably going to win the battle or probably going to lose the battle. One also asks questions like, under what conditions, with what barometric variations is this supposed to have happened? I think that these big questions and big answers are very dangerous to ask and to give answers to because almost everybody will agree that the numbers are not always correct, that the assumptions are very critical and can always be argued about as can the numbers. The exact answers such as days, weeks, of duration of the war, or the change or the probability of winning a war or losing it is very difficult to get from models. It is also very problematic that definite confidence levels can be given for these types of answers.

I would like to make one proposition, which certainly comes from, to say the least, the very much shorter experience we have had in Germany with modeling at the theater level. This proposition is that one should use models at the theater level, and certainly more below that level, so to look more at the functioning of what was described earlier as the phenomenology of combat, and the complexity of the interaction of forces as it develops on the battlefield. How does the model battlefield look after five years, after ten years? This is a very important question for all planners. They have to bring new weapons, new tactics, and new doctrines into this environment for which the model war of 1973 was or is perhaps a better model than a few other models, but with its own shortcomings too. I would like to propose to you that we think about theater level gaming and below, more in terms of looking at the rules, the interdependencies of the battlefield, to be able to better understand the possible effects of new technologies, the possible effects of force structure changes. Simulations can really produce very useful data, which are needed for further analyses, and sometimes for planning itself. If you only look at the problem of defining probabilities in the sense of how often does it happen that a certain weapon system is confronted with another weapon system, these are certainly questions an equipment planner has to answer. If there are no models like the Model War of October 1973, the planner has a difficult time answering such questions. I feel that the specific strength of simulations in models is in giving insights into these interdependencies on the battlefield, which can produce frequencies of distribution and from that perhaps new probabilities that are needed to define needs and future developments.

Dr. Goad: I come from the SHAPE Technical Center where we conduct analyses of force capability for SACEUR and also do some work for the NATO community at large. I'd like to address two points that have been raised by several speakers this morning, one which I think Andrew Marshall brought out indirectly about the use of models — what can they do for the decision maker? He concluded, I think, that they were perhaps not of very much use.

The second point I'd like to address, albeit briefly, is what we as analysts ought to do, or could do, concerning communicating the results of our analyses to decision makers. There's not much point in doing all this work if after we've done it we don't have the means at our disposal to address the decision makers who we're trying to influence.

Regarding the first point — what are the models useful for — the previous three panel members have all addressed various aspects of this problem. I'd like to just bring out two points that are perhaps obvious but may not have been made with the force that I think they ought to be made. First what the models cannot do — they cannot predict in absolute terms the outcome of combat. Dr. Linnenkamp made this point. We do not know, given the models and the technology we have at the moment, how long the war will last in Central Europe, we do not know what the casualties will be, and so forth. However, having said that, what we can do, and I think it's useful, is to identify trends in force capability. We can, for example, perform analyses of force capabilities at a given time, say, now, and we could (and I think we should.) compare those estimates with what the answers might be with the

same models in, say, 1982. If we were to do this, we would be able to address quantitatively a problem that greatly concerns senior military people as it does our budgeters and politicians, and that is, just how big is the Warsaw Pact getting? How capable is the Warsaw Pact soldier that we postulate? Is, in fact, Warsaw Pact capability outstripping that of NATO? Yes, or no? Qualitatively, many of our senior military people would think that this were the case, and indeed it may be the case. But, I think, as analysts, we should do the work — find out whether it is true and address particularly those aspects of force capability that contribute most to the increase in the Warsaw Pact threat, if indeed the offensive capability of the Warsaw Pact is increasing.

Another problem area in which the models can be used is the comparison of alternatives. One obvious example is in the comparison of alternative force deployments. There's been much talk about the maldeployment problem, what should we do about it, what are the alternatives, and what are the plusses and minuses associated with those different alternatives. So, I think the models do have something to say about comparative force capabilities, with or without the limitations concerning things that are not considered such as command and control, which is the most important, and training, morale, leadership, and all the rest of it.

I'd like to also address as my second point, which I think is probably even more important than the first because the first has been talked about at length for many years now. I'd like to talk about the communication of the results to the decision maker. All I can suggest is a somewhat indirect approach concerned with trying to bridge what I like to think of as the credibility gap. I think that the decision maker is entitled to look askance at the models that we now have, as perhaps Andy Marshall has done, when he sees the results that the U.S. produced with their models and compares them with perhaps the results that the FRG produced, and then in turn with those that the UK produced. They are apparently all different.

An associated problem lies within the NATO community. For example, when the UK writes a report concerning force capabilities, which is obviously a UK position, and someone tries to put it across in one form or another within the NATO environment that tends to get dismissed perhaps more simply than it ought to be as just the UK position. There's no reconciliation with the U.S. position or the FRG position because different models have been used with different assumptions.

Under the auspices of the MCSSG there's been an attempt to compare the effects of three different methodologies on the same theater situation — same forces, same Warsaw Pact threat — to identify the extent to which the models are different, and to identify the reasons for those differences, and, hopefully, to reconcile those differences. This involved a lot of work and turned up a lot of interesting results. One result was that the models turned out to be not as different as they appear. The main difference appears to be in what I will call the "rate" at which comparable events on the battlefield appear to occur. Some models go faster than others, but the trade-offs, for example, between the rate at which casualties are suffered for ground given up, and so forth, are, in fact, very similar. I think it's interesting, broadly speaking, that the three models tended to portray the same battlefield development. It was as if each of the nations had taken a film of the battlefield development and were projecting it at different speeds. This is significant because it says that, although the three nations used very different models with very different structures, the only real difference between the models is the time factor, the rate at which comparable events appear to occur.

The very real rate difference is important because there are many time-related factors, particularly those concerned with the introduction of U.S. reinforcements into Europe, which are clearly critical when it comes to decisions about the actual time it takes for various events to happen. I think that one is justified, for example, in asking whether the U.S. should spend a lot of time and effort trying to increase or improve the rate or the timeliness with which U.S. reinforcements arrive in Europe. Obviously, if it's a very short war we're talking about, bringing U.S. reinforcements in two days earlier is not going to help much. If we're talking about a long war, it's not going to make too much difference either. The only case in which it's going to matter is when the arrival of the reinforcements will presumably change the outcome of the conflict. In other words, the message we should be stressing is that models cannot predict the outcome of combat in absolute terms. I'm not really sure how we should do this, that is, how one can determine the proper arrival time for reinforcements, but it does address this credibility problem and it is one of the reasons that I think some people, some decision makers, look askance at the value of the models that we're all using.

Dr. Kupper: Before the question period, I'd like just to make some points that were provided to me that reflect some concerns with some statements of this morning.

First, models tend to be accepted — at least in the presentations that were given — the point was made that models are accepted and that the answers are unquestioned. Those of us that work in the field know that's not true, but that came across as a message in three of the talks this morning.

Now, relative to PRM-10, the use of the measures of the capability indices, these were very carefully annotated with comments and supplied for example, with caveats. Unfortunately, the comments that were done within the services, were deleted when the posi-

tion paper went over to NSC. That deletion was really glossed over and that was very important. The SAGA conference on PRM-10 decided the MOCs for contingency assessment. These were agreed to after a lot of very tough discussions if you will. Unfortunately, they were lost.

The second point I want to make relates to hierarchical models. Apparently no one really touched on the judgmental aspects of what is passed from one model to the next within a hierarchy. Again, we're probably making assumptions here and that could have been touched on too lightly.

Another point that was not mentioned but which perhaps should be looked at is the fact that the range of uncertainty should be displayed for decision makers. This ties in with the point I think several of the speakers have already made, that models can't predict in absolute terms so to speak, but by the same token they can show directions, trends. You know, again, this is part of this understanding the phenomena that I mentioned.

We're now open for questions please.

Mr. Louer: I'm afraid that maybe some of the things I said this morning may have not been conveyed properly. Remember, my presentation was in three parts: number one, what the Army force planning process has to do, what's necessary: number two, how this translates into requirements for theater models: and number three, how theater models are currently being used in this exercise.

Number one, the Army force planning process must determine absolute estimates of how many divisions and how many forces it has to have. I used those words — that it must determine absolute estimates. The Army has to make that decision. Now, in the third part I referred to certain models that were being used in support of these decision processes. I did not mean to convey that they were using absolute results in those models. Those models are being used in comparative modes. They are looking at the ranges of estimates, and so I just wanted to clarify that because this issue of absolutes did come up.

Also, about the other issue that came up relating to the complex nature of theater warfare, I think I tried to emphasize that, too, throughout. Depending on the forces, the command control and communications, and intelligence, it's a very complex process, which generates many different combat situations. Unfortunately, up to this point, we're only able to comprehend where the FEBA is at the end of the battle and how great the losses, and we start using these numbers. That's bad. That is too simple an answer for this complex process. We do need better ways of measuring the alternatives.

Dr. Kapper: The points that are being made in here I think are all well taken and as Larry (Low) mentioned, one of the things we hope here is that people come across with some new directions and I think these are some. Next question.

Dr. Huber: Actually I have a kind of comment. I would be very interested in hearing the panel's opinion. Often when decision makers, or let's say users, judge a model, it seems to me they do not have a very precise conception of what that model really means. The pie chart that Dr. Kapper illustrated shows how little of the total decision makeup process depends on models. Now I would like to make a statement, if I may. Certainly every decision maker, every user, every general, has some kind of model that he uses for his decision making. The main difference, however, between his model and our models is that his model is very often not very well structured, while ours is at least structured.

Dr. Kapper: That's an excellent point.

Dr. Huber: What we try to produce by these models is answers to questions like, "what happens if?" We do not want to predict. This is one comment that I'd like to make to Rex Goad's statement that we probably ought to be devoting more time to the time factor. If an analyst came around and told me that he has a model that predicts the time of a war, I'd be very suspicious because it would be better than reality. A model is the better the more it reflects reality, but I would question a model that could do better than reality because reality is just one sample out of an invariable number of possible developments, and because reality is highly stochastic. So, I'd be very suspicious if somebody said the war is going to last six days. It could be ninety, it could be two.

Dr. Goad: Yes, what I didn't address was the formula in which we ought to be providing these predictive answers. Obviously, it would have to be stochastic, because real life is fraught with uncertainty, and we should be looking for the distribution of the results. I think it's legitimate to suggest that the force planning process really ought to be performed within an environment in which the tails of the distribution should be taken into account. What would happen if the 10% worst cases were to occur? Presumably, we have to cut off the tails and say, "well, we can't insure ourselves against the very worst cases that can happen, but we ought to be taking account of 95% of the distribution."

Certainly, nothing I said was intended to suggest that we should be using deterministic models, far from it.

Dr. Huber: I only meant to say that to evaluate a model by whether it can predict the time of a war or not is the wrong proposition.

Dr. Kapper: Question, Frank?

Mr. Hoeber: At the risk of sounding like the dissenter on everything, I like to say that I'm somewhat concerned that we don't lapse too much into hubris about what modeling and, more than that, what analysts can do. It strikes me that often the decision maker's larger model — larger structure — is not necessarily unstructured, if he's a good decision

maker, and we have almost as much variety among decision makers as among analysts. I think what we are, or should be, talking about is whether our models or structures are well defined. I would expect, pursuant to your pie chart, Frank, that we never structure as completely as a decision maker until we become a decision maker ourselves, but we ought to structure what we *can* structure very well. More importantly, we ought to articulate it, and this is perhaps the start of rendering a service to users.

Dr. Kapper: Other questions?

Mr. Asbed: I'm going to address my question to Mr. Hoeber. My question is about the importance and viability of sensitivity analyses in deterministic models, where there is higher and higher resolution introduced, or in stochastic models, where you have to use Monte Carlo techniques. Would you comment, please?

Mr. Hoeber: Well, I'm not sure that there's any definitive answer to that, although others here may want to try. But I'm willing to comment on the utility of sensitivity analyses in two different kinds of models, the deterministic and the stochastic where presumably Monte Carlo type results are being used. Now, we have great problems, which have been mentioned several times, with the deterministic models. The trouble with deterministic models is they're the kind we know how to do. I see, however, and there may be others who want to comment on this; I see no conflict if the stochastic model you are using, which, for the sake of argument, is a strict Monte Carlo procedure, and is a good one which is giving you reliable results in your estimation. And I see no conflict if you are able to replicate, or can afford to replicate, enough so that the range of the results can be usefully narrowed. Then it seems to me that sensitivity tests still accomplish their purpose.

Perhaps, as a more general statement, I should say, as others here have, that you cannot really simulate the real world. If you get too complete you've made it too deterministic, and you've only got one sample out of the real world. What you are really trying to understand is the process, and it seems to me that sensitivity tests help you know when you're getting an understanding of the process. You can't test the sensitivity of all the combinations when you have a large number of variables and a large amount of interaction among them. You use all sorts of techniques, "pre-processors," as Jasper (Welch) called them, to try to pinpoint the interesting cases, to eliminate many of the cases, to isolate those cases at the extreme where you're succeeding or failing, as Jasper's P_k example, and then test those further. It seems to me what we ought to be focusing on, and, if talking about sensitivity tests sound like you're overmeasuring your results for a given data base and assumptions, maybe we should just use the term "understanding the process." I believe we've got a pretty good consensus that we're not trying to get answers in any absolute or outcome sense. Maybe that is terribly disappointing to us with our pride in craft, but does any brave man want to give an example to the contrary?

Dr. Kapper: More questions?

Mr. Low: I'd just like to ask one question of Norig (Asbed). Were you addressing in any way, Norig, the trade-off between degree of resolution and going deterministic or stochastic?

Mr. Asbed: No, actually I was saying that when you have a deterministic model and you have a very high degree of resolution a sensitivity analysis does not seem to me to be very helpful. There is too much detail in the model that's going to mask the item you're looking into.

Mr. Low: I see, yes, I think that was the point. I think that's a good point. You may have an over-determined model if you have many, many variables.

Mr. Asbed: That's right.

Dr. Kapper: Are there any other questions?

Dr. Dobieski: I have a comment to make about theater level modeling. I don't know what the requirements are. In other words, I hear some people saying they're not satisfied with the models, and some saying they are satisfied with the models. I would like to see a hard requirements document generated for theater level modeling. I'm not talking about a wish list, I'm talking about a hard testable requirements document, because I don't know how to measure how good the models we have are unless I really know what the requirements of the users are, and I haven't seen such a document. I see attempts at a good first cut. I was wondering if other people share my opinion that this would be a good place to start?

Dr. Bonder: Maybe I could comment on Dr. Dobieski's comment. I think the user community is schizoid. On one hand I hear them say the models don't capture all my detail, and Andy says they don't have the leadership and the intelligence and all the detail in there that one wants. Then I hear somebody else say, "Oh, but they take the long draw, and I've got to spend four people on it, and they've got $150,000 a year to spend to run it" — they've got billions of dollars they're investing in resources. I think it's a schizoid user community. They don't know what they want.

Dr. Kapper: There are diverse opinions, that's true. Are there any other questions or points?

Dr. Bracken: I'd like to comment on what appears to be somewhat of a logical disconnect in Andy Marshall and Rex Goad's comments. This morning, Andy said models are not descriptive of any military balance, and Rex said one thing models cannot do is predict outcome in absolute terms. Frank Kapper was talking about the JCS/JSOP measurement of risk, which is sometimes done when a number of divisions are required to take a FEBA from a very bad place and put it on a border. That's a very absolute number. It

comes from an absolute reading of models in many cases. Rex said that one can perhaps identify trends. Well, if you have two absolute numbers and you want to find an interesting trend, you have to somehow believe the absolute differences between these absolute numbers in order to identify a trend.

Now, I'm not comfortable recommending that we believe the outcomes of models of absolute numbers, but I think we ought to recognize what it means when we say that we can't believe these things.

Dr. Kapper: Do you want to comment Rex?

Dr. Good: Okay, what you can't do in absolute terms you can in comparative terms. In other words, you've got a graph with no measurement on the Y scale.

Dr. Bracken: That's exactly my point.

Dr. Good: You just don't define your units period, and then you're home — home free in theory.

Dr. Kapper: With respect to your point Jerry, let me just say that instead of saying it is ten, you can say it's somewhere between this and this and then you avoid that absolute nature, in a mathematical sense. I would say we play the range game and in practical terms we really often deal with the models as though they were absolute. I think that's a good point. By the same token, it's a caveat, which I mentioned that applies to everybody, JCS included.

Session II — Status of Theater-Level Simulation Models, Present and Future, and Problems of Model Structure

Opening Remarks

DR. JEROME BRACKEN
Institute for Defense Analyses

Mr. Low: We'll now proceed to session two. Our session chairman will be Jerry Bracken from IDA.

Dr. Bracken: I'm going to start off with some specifics that may be useful. The study that I want to mention very briefly involves comparing alternative weapon system mixes. The study could be considered as not involving absolute numbers, so we avoid that problem for the time being.

We performed this study for J-5. The title is the "Cost Effectiveness Study of NATO Force Improvements," a WSEG report published in October 1975. This report isn't very widely distributed, but we were able to publish a summary in the Journal of Defense Research and the citation is shown. The Journal of Defense Research is a classified Journal and the study was classified, but my remarks are taken from the unclassified portions.

What we did was look at a number of NATO force improvements. This amounted to a five billion dollar ten year cost force improvement package. That's a half a billion dollars a year on the average. We were putting together force mixes consisting of increasing shelters for aircraft, substituting TOWS for recoilless rifles in combat units, increasing M-60's, or TOWS in combat units, increasing weapons in replacement pools, increasing the stockpile of Maverick air-to-surface missiles, and increasing the surface-to-air systems.

We used allocated indirect costs in costing out our increases in unit weapon systems. We put together a number of packages, a hundred or so different alternatives, and we narrowed these down to about ten alternatives of five billion dollar force structures. There were significant differences in war outcomes, relative outcomes if you like, between these kinds of force improvements. There were interesting nonlinearities. For example, if our measure of effectiveness was the position of the FEBA at T+30, a replacement tank that's used on D+30 wasn't as effective as a unit tank added on D+1. So the packages were widely different, but also quite complimentary.

We used the IDAGAM model in the study. We used the SAGA data base that had been used in a previous study. One of our recommendations in the study was that this analysis be done again with one of the other theater level models to increase its credibility if the same conclusions were obtained. But, again, let me emphasize that the model outcomes were very sensitive to this range of changes.

Now, I'd like to ask Seth Bonder to give his talk on theater level models. Seth Bonder is the President of Vector Research and has been in the military operations research business a long time.

8 — Theater-Level Models

DR. SETH BONDER
Vector Research Inc.

Vector theater model evolution and development.

Dr. Bonder: I'll talk a little bit about the Vector series of models. I have one slide on all of them, plus some trends, and some notes about the schizophrenia mentioned today that I want to talk about later.

The VECTOR series models really started with Glenn Kent back in 1972 when he had three organizations in competition — IDA, Lulejian, and Vector — building different theater level models. The charge to us at the time was to build a theater level model without the use of firepower scores, neither for attrition, nor movement, and to somehow decompose what was called the Army in the big ATLAS game into fundamental weapon system types.

The approach that we took at the time is shown on Slide 8-1. We decided we would try to simulate, if

SLIDE 8-1 VECTOR MODES PROJECT OBJECTIVE

DEVELOP A THEATER-LEVEL ENGAGEMENT MODEL WHICH:

- USES DETAILED MODELS OF THE PHYSICAL AND BEHAVIORAL PROCESSES IN A CAMPAIGN
  - DOES NOT USE THE "FIREPOWER SCORE" FORCE RATIO CONCEPT OF ATTRITION BUT ONE WHICH REFLECTS INTERNAL DYNAMICS OF THE COMBAT ACTIVITY
  - DISAGGREGATES THE ARMY SO THAT ATTRITION OF PERSONNEL AND INDIVIDUAL WEAPON SYSTEM TYPES CAN BE DETERMINED
  - DRIVES MOVEMENT OF THE FEBA BY MEANS OTHER THAN THE "FIREPOWER SCORE" FORCE RATIO CONCEPT
  - CONSIDERS ALL RELEVANT COMBAT AND NON-COMBAT PROCESSES
- PROVIDES THE USER WITH FLEXIBILITY TO PORTRAY ALTERNATIVE FORCE STRUCTURES, DOCTRINES, WEAPON SYSTEMS, AND TACTICAL DECISION MAKING BEHAVIOR

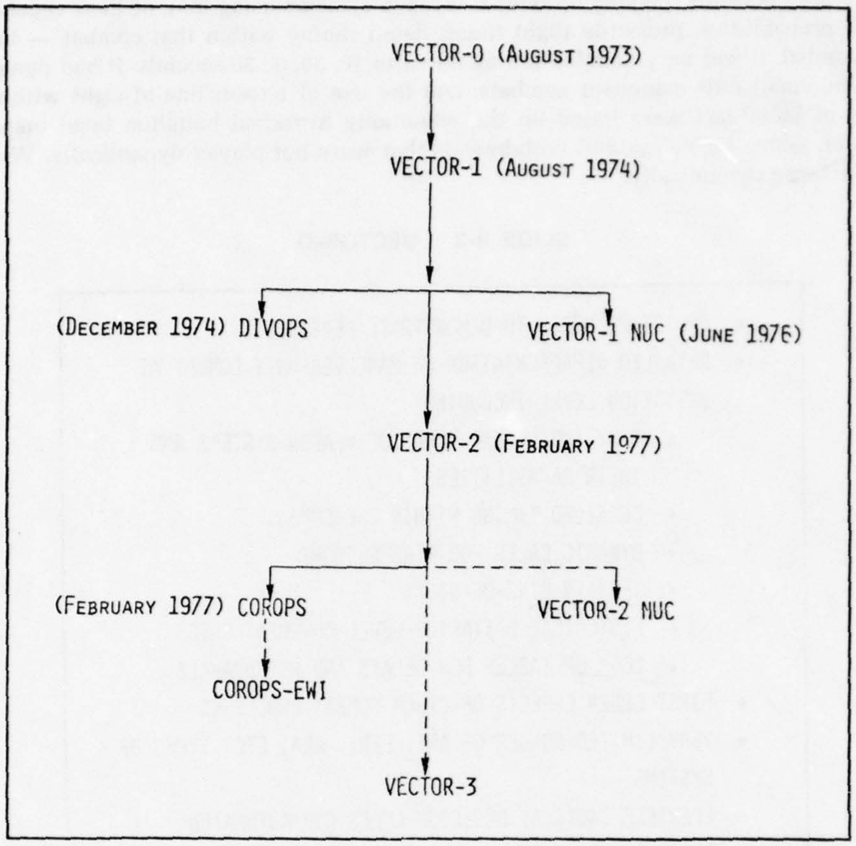
DEVELOPMENTAL APPROACH

DESCRIBE IN DETAIL

- TYPES OF SYSTEMS
- ORGANIZATION FOR COMBAT
- BATTLEFIELD REPRESENTATION
- ATTRITION PROCESSES
- NON-COMBAT PROCESSES (C^3, INTELLIGENCE, MOVEMENT, ETC.)
- TACTICS AND TACTICAL BEHAVIOR

DESCRIBE IN TERMS OF PHYSICAL AND BEHAVIORAL VARIABLES AND THEIR DYNAMIC INTERACTIONS

SLIDE 8-2 VECTOR MODELS- -LINEAGE



you like, the war in some sort of detail. Detail in types of systems, formulization for combat, the battlefield representation, the various attrition processes, many of the noncombat processes like intelligence, electronic warfare, perhaps supplies, and the tactics and tactical behavior, without the use of firepower scores. Now, clearly to do that was not a one-year task, and we agreed on a sequential development program. I'd like to show you a little bit of the genealogy of that program, which involves several models (Slide 8-2). I'll discuss each of these very briefly.

VECTOR 0 was essentially the prototype to see if we could do anything intelligent without the use of firepower scores. It turned out that it was feasible and I found it very easy to do. VECTOR 1, which was developed in 1974 gave rise to two other sub-pieces, something called DIVOPS, a division level spinoff of VECTOR 1 that was reviewed by the Army in a number of studies, and something called VECTOR 1-NUC, which was an integration of a nuclear salvo model with the VECTOR 1 conventional war model for use in some combined conventional nuclear studies. VECTOR 2 was just completed in February 1977. I'd like to say it is an enrichment of VECTOR 1, but it really is a brand new model, a whole new structure, which I'll talk about briefly.

There is a spinoff from VECTOR 2, which we call COROPS. By just pushing one switch you now have a corps level model, or division, if you like, because the model computes by sector and that way you can spin off one whole sector for a corps. We had planned to do a VECTOR 2-NUC model, analogous to the VECTOR 1-NUC, and there are some proposals outstanding for that. We have something called COROPS-EWI (the dotted lines indicate those tasks are not under way now), which is an attempt to integrate some electronic warfare, intelligence, and communications methodology we've just finished developing for CAA into the structure of the VECTOR 2. But of course it is only a slice of VECTOR 2, so you will get essentially many of the noncombat processes well integrated with the heavily combat-oriented Vector series.

Finally, there is a dream in my eye called VECTOR 3, which involves the nonintegral FEBA. This is not in any of the models at this point. The dotted lines are where we'd like to go, and the solid lines are things that are completed now and available for use.

Let me talk a little bit about each of these. As I pointed out, VECTOR 0 was the prototype model to demonstrate the feasibility of modeling theater warfare without firepower scores (Slide 8-3). It had a fairly detailed representation of the maneuver unit combat, especially down at the battalion level, mathematically at least, with the preprocessors, playing individual weapon systems using may of their capabilities such as hit probabilities, kill probabilities, projectile flight times, detail timing within that combat — that is, within the maneuver unit combat, it had very detailed timing down to 10, 20, or 30 seconds. It had dynamic calls for fire support during the small unit maneuver combats, and the use of terrain line-of-sight within those combats. The movements of battalions were based on the essentially historical battalion level maneuver rates, and there were look-up tables for delays and withdrawals that were not played dynamically. We only played the assault on each defense dynamically.

SLIDE 8-3 VECTOR-O

- PROTOTYPE MODEL TO DEMONSTRATE FEASIBILITY
- DETAILED REPRESENTATION OF MANEUVER UNIT COMBAT AT BATTALION LEVEL INCLUDING
  - MANY INDIVIDUAL TYPES OF WEAPON SYSTEMS AND THEIR CAPABILITIES
  - DETAILED TIMING WITHIN THE COMBAT
  - DYNAMIC CALLS FOR FIRE SUPPORT
  - TERRAIN LINE-OF-SIGHT
  - HISTORICAL BATTALION-LEVEL MOVEMENT RATES
  - LOOK UP TABLES FOR DELAYS AND WITHDRAWALS
- FIRST ORDER EFFECTS OF OTHER COMBAT PROCESSES
- VERY LIMITED NUMBER OF ARTILLERY, ADA, ETC. TYPES OF SYSTEMS
- FLEXIBLE TACTICAL DECISION RULES FOR AUTOMATED DECISION MAKING
- IMPLICIT REAR AREA
- IMPLICIT C^2, INTELLIGENCE, AND COMMUNICATIONS
- AGGREGATED OVERALL TIMING

There were first order effects of other combat processes, a very limited number of artillery, one type, I think, and the same with our defense. There were flexible decision rules for automated decision making inside the model. Essentially you could play any type of behavior you would like to play.

There was no rear area in VECTOR 0. It had very implicit command control, intelligence, and communications that were essentially built into the data base for intelligence, and very aggregated overall timing. That is, the main clock of the model was a day, or if you want you could make it twelve hours, but somewhere between the beginning of that day and the end of that day the battles take place.

VECTOR 1 (Slide 8-4) was really a slight enrichment of VECTOR 0, without a great many changes. It was the first production version of the VECTOR series. It had more maneuvering of weapons than in VECTOR 0, more artillery and air defense artillery, but a limited number — four artillery, six air defense artillery. We also included more terrain features in VECTOR 1, mine fields, and the use of aircraft shelters. We developed a preprocessor which would automatically hook into the OSD data base, theoretically both sides of the data base, that is, the order of battle data to call up all the NATO forces and all the so-called Pact forces and organize them on the front, as well as the weapons data, although there are no weapons data, as such, in the OSD data base for the model, but it automatically will hook into it and call for whatever is there.

VECTOR 1, as I said, was completed in the late summer or early fall of 1974 and is now running on a number of computers as I've shown you here. The one we're using at VRI now is the AMDAHL 460/V6, which is a Japanese version of the IBM machine. I think it is also used at a number of agencies.

SLIDE 8-4 VECTOR-1

- FIRST PRODUCTION VERSION
- MORE MANEUVER UNIT WEAPONS
- MORE ARTILLERY, ADA, ETC. TYPES OF WEAPONS
- TERRAIN AND TERRAIN FEATURES EXTENDED
- MINEFIELDS
- AIRCRAFT SHELTERS
- PREPROCESSOR USING OSD DATA BASE
- RUNNING ON:
  - IBM 370/168 AND AMDAHL 460/V6 AT VRI
  - UNIVAC 1108 AT CAA
  - CDC 6400 AT WSEG
  - HIS 6180 AT AFSA
  - HIS 6080 AT CCTC

SLIDE 8-5 DIVOPS

- A NON-PLAYER ANALYTIC REPRESENTATION OF COMBINED-ARMS ACTIVITY AT DIVISION LEVEL
- DRAWS HEAVILY FROM VECTOR-1 THEATER-LEVEL MODEL
- INCLUDES ENRICHMENTS NOT PRESENT IN VECTOR-1
  - SIMPLIFIED REPRESENTATION OF REAR AREA
  - TARGET ACQUISITION FOR FIRE SUPPORT
  - MORE DETAILED TIMING OF EVENTS
- RUNNING ON:
  - IBM 370/168 AND AMDAHL 460/V6 AT VRI
  - UNIVAC 1108 AT TRASANA

I'd like to say a little bit about DIVOPS (Slide 8-5), which is a spin-off of VECTOR 1. We actually cut out a lot of pieces of VECTOR 1 to use for a study on the Army's requirements for Air Force close air support. The only reason I'm showing it to you here is because it was part of the evolution. In other words, we added to the division level slice of VECTOR 1 part of a rear area in a very simplified form, and we explicitly added target acquisition in the rear area, that is, imagery sensors looking for targets in the rear area for fire support. Another thing that is interesting here is that we had more detailed timing of events rather than a day. It was down to 30-second computation steps in the division model. DIVOPS is running on a number of computers both at TRASANA and at VRI.

About a year and a half ago, in looking at the NATO capability to stop a Soviet granular attack for DNA, we added to VECTOR 1 a nuclear salvo model called UNICORN, which was mentioned earlier. That is, we literally made UNICORN a module of VECTOR 1 and called the resulting model VECTOR 1-NUC (Slide 8-6).

SLIDE 8-6 VECTOR-1 NUC

- REPRESENTS THE COMBINED TWO-SIDED/CONVENTIONAL TACTICAL NUCLEAR CAMPAIGN BY INCORPORATING A DETAILED NUCLEAR FIRE ASSESSMENT MODEL INTO VECTOR-1
- ALLOCATES NUCLEAR FIRES TO CONVENTIONAL TARGETS AND TNF ELEMENTS CONSIDERING MANY WEAPON TARGET CHARACTERISTICS
- ASSESSES DAMAGE (RADIATION, BLAST, THERMAL) TO INDIVIDUAL TARGETS FOR ANY INDIRECT FIRE WEAPON AND TACTICAL AIR
- PERMITS ASSESSMENT OF EFFECTS NUCLEAR FIRES HAVE ON SUBSEQUENT CONVENTIONAL COMBAT RESULTS
- RUNNING ON AMDAHL 460/V6 AT VRI

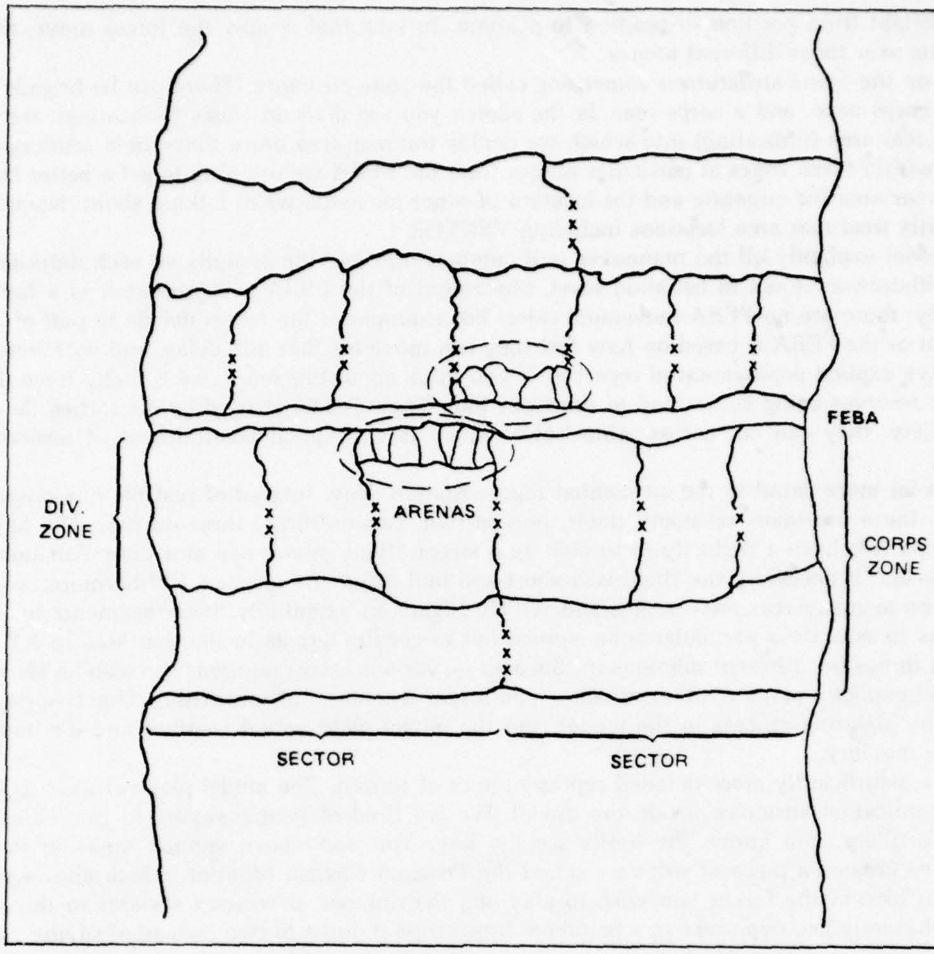
The model plays combined and conventional nuclear warfare, but the nuclear portion goes in salvo, and not in continual fires. You can stop the conventional war at any time you like, plug in the nuclear portion, look at the resulting forces, and then continue to play the conventional game. It plays essentially all of the nuclear aspects of UNICORN and all the conventional aspects of VECTOR 1. The results of the DNA study in which we used VECTOR 1-NUC have been briefed literally all over Europe and the United States.

SLIDE 8-7 VECTOR-2 (COROPS)

- INTELLIGENCE/TARGET ACQUISITION PROCESSES ADDED
- COMMUNICATIONS PROCESS ADDED
- COMMAND AND CONTROL PROCESS BROADENED
- TO PORTRAY REALISTICALLY THESE PROCESSES, THREE NEW STRUCTURES WERE DEVELOPED:
  - ESSENTIALLY, CONTINUOUS REPRESENTATION OF TIME
  - EXPLICIT COMMAND LEVELS FROM BATTALION THRU THEATER
  - BATTLEFIELD REPRESENTATION
    - COMBAT ARENAS
    - ZONES
    - SECTORS
- ALL MANEUVER UNIT ACTIVITIES TREATED DYNAMICALLY
- MOVEMENT OF FEBA EXPLICITLY REPRESENTED AS A FUNCTION OF COMBAT ACTIVITY
- EXPLICIT EMPLOYMENT OF RESERVES
- MORE DETAILED REPRESENTATION OF:
  - AIR COMBAT
  - CAS MISSIONS
  - INTERDICTION MISSIONS
  - ESCORT MISSION
  - WEATHER EFFECTS
  - MORE DETAILED REPRESENTATION OF TERRAIN
  - DECISION AND COMMUNICATION LAGS
- PROGRAM CHANGE MONITOR
  - INCREASE OR DECREASE DIMENSIONALITY OF VARIABLES
  - ADDITION OR DELETION OF SUBSCRIPTS
  - ADDITION OR DELETION OF VARIABLES
- POSTPROCESSOR
- RUNNING ON
  - AMDAHL 460/VE AT VPI
  - HIS 6080 AT CCTC
  - IBM 360/65 IN PRIVATE INDUSTRY
- DOES NOT EXPLICITLY TREAT:
  - NON-INTEGRAL FEBA SITUATIONS
  - TACTICAL NUCLEAR WARFARE
  - ELECTRONIC WARFARE

Finally, we have VECTOR 2, and also COROPS (Slide 8-7). COROPS is exactly a slice of VECTOR 2, unlike DIVOPS for which we had to cut all the pieces from VECTOR 1 and tie them all up. There's nothing to do with COROPS, you just push a button, set "S" equal to one second, and you have a corps level model. Okay, we added to the concepts in VECTOR 1. We thought we were going to enrich it to make a VECTOR 2. The purpose was to add intelligence and command, control, and communications to the process. It turns out we had to build a whole new model. When I say intelligence, I mean estimates of the order of battle. The model literally takes detection data — acquisitions — and makes continual estimates of the order of battle of the forces opposing you, but it also uses a lot of the target detection information for fire support target acquisitions, and it will allow you to play first order effects of that in the sense of using imagery sensors, ELINT but not all of COMINT and HUMINT. Anything you can really talk about as a way of sensing targets.

SLIDE 8-8 BATTLEFIELD REPRESENTATION



Now, we have a communications process added in there, essentially a log scale queuing network for communications of messages. The command control processes were broadened extensively from VECTOR 1 where they were very implicit.

To do all of that we had to add a whole bunch of new structures to the models — essentially we went to continuous representation of time. No longer does the war begin at 8:00 in the morning and at 8 o'clock the next morning you come back to see what happened. Instead, you literally follow through in real clock time. It's continual based: There are a whole set of clocks, five or six, set for the appropriate timing of the events so that the small unit combats may compute on 30-second bases, and the target acquisitions in the area may compute on a 15-minute basis or a 5-minute basis.

We have an explicit command hierarchy, both for the NATO nations and the Warsaw Pact, and you don't have to play them symmetrically in any way. As suggested by Andy (Marshall) you can play your tactics differently, you can play your command structure differently.

We added a significant amount of detail in the battlefield representation using three concepts called combat arenas, zones, and, of course, sectors. Sectors consist of one or two corps, if we're thinking about the European front. The concepts of zones and arenas can be shown by a little sketch (Slide 8-8). The sketch shows a split FEBA, and if you had a line FEBA there are two forces that are separated by some distance which is what in reality is at the single line. The arena concept involves a maneuver unit, either battalions or brigades, or whatever you'd like to push in there, that do fronting. If you're talking about command and control, you've got to talk about fronting decisions — who opposes whom. The arena concept is essentially a map of the total terrain and areas into which units deploy for front line combat and in which they engage and move from arena to arena to arena.

The whole European theater is divided into arenas, roughly I would say into areas of 6 kilometers on a front and 15 kilometers deep depending on terrain features. There are objectives within the arenas for which

the units will fight from position to position to position. In fact, that is how the forces move, they move in maneuver units over those different arenas.

Overlaid on the arena structure is something called the zone structure. There can be brigade zones, division zones, a corps zone, and a corps rear. In the sketch you see division zones (indicating), the corps zones and the corps rear area (indicating) into which we deploy the rear area units. Since field artillery and the like are deployed within these zones at particular ranges from the FEBA we're trying to get a better handle on the location of a rear area for targeting and for location of other elements we're talking about. None of the other models currently treat rear area locations including VECTOR 1.

We do model explicitly all the maneuver unit combats, not just the assaults on each defense, but all the delays and withdrawals down to battalion level. Movement of the FEBA is represented as a function of the combat activity; there are no FEBA movement rates. For example, if the forces decide to pull off a particular hill, movement of the FEBA is based on how fast they can move off that hill, delay, and withdraw.

We do have explicit deployment of reserves. If you think about the zones, we actually trace the XY coordinates of the reserves being committed to the front line. They can be attrited by air sorties they can be attrited by artillery, they can cut across mine fields. There is an explicit commitment of reserves and their movements.

We added a lot more detail to the air combat representation. Now instead of just the one-on-one duels we used to have, there are many-on-many duels, one-on-two, two-on-three, three-on-four. We have close air support missions. We have a flight fly in to pick up a target, flying passes one at a time. Attrition is assessed after each pass and if necessary the flight will abort and pull off of the mission. Furthermore, we have interdiction missions to many rear area targets and we fly flights on essentially linear segments to their targets. They can mass to saturate a particular zone, spread out to get the targets in the rear area in XY coordinates and figure out things for different missions in this area — various escort missions can also be accommodated.

The model explicity plays weather effects — both real weather and forecasting. Our five day forecast is entered into the planning process in the model, and the model plays actual weather and it's impact on both ground and air mobility.

We have a significantly more detailed representation of terrain. The model plays officers in the decision and the communication structure inside the model. We got tired of people saying to us, "Chief, you don't have enough artillery, you know, the limits are too low. You don't have enough tanks in the maneuver units." We have created a piece of software called the Program Change Monitor, which allows you to scope the size, if you like, of the forces you wish to play and the number of weapon systems in them. So, if you have a study that only has two tanks in a battalion, fine, scope it down to two instead of saying "let the number of tanks be two" and save core and running time. This Program Change Monitor allows you to increase or decrease the dimensionality of variables, and to change the subscripts on variables, which alters their dependency throughout the model, and, in fact, it can help you add and delete variables from the model. It will show you (you've got to do some coding yet), where all the interconnects are so that program changes are relatively straightforward — not easy, but relatively straightforward.

There is a postprocessor that stores essentially the total combat results and that will print out all the tables and graphs that you ask it for. The model is now running on the AMDAHL 460/V6 at the University of Michigan. The model is partly being converted to operate at CCTC on their 6080, and we have a contract to put it on a computer for one of the private companies in the United States.

The model does not explicitly treat nonintegral FEBA situations. We do not allow air drops, and that's a word we've been using in the community for about three years or four years. That is, the Blue and the Red must always be separated. I don't care what the shape of the FEBA looks like, but in this model I can't have Red drop behind Blue. It does not treat tactical nuclear warfare explicitly, and it does not treat electronic warfare explicitly. But, you can treat that implicitly with the data base, just as you can do nonintegral FEBA implicitly if you're clever, but the model wasn't designed to do that.

Okay, any questions on the VECTOR series?

Mr. Asbed: How much backup do you need for this?

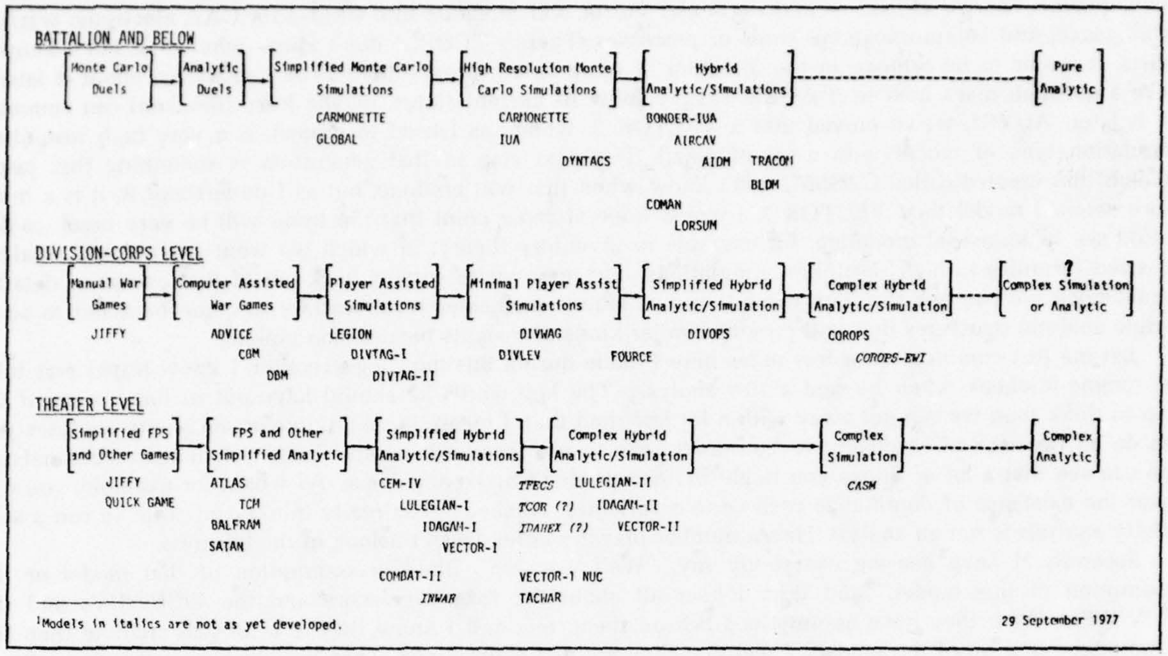
Dr. Bonder: Lots — a lot of data. I want to say 360,000 pieces of data or something like that.

Mr. Asbed: Where do you get them?

Dr. Bonder: From the military, from exercises, from laboratories. They are all measurable variables. Every variable in that model is a measurable quantity, and I will talk a bit about that in a minute.

But first let me now talk about our other models. I put together this chart (Slide 8-9) about a year ago, and modified it a bit last night. It's intended to show you some trends in theater level modeling, but I want to compare it to the battalion level model and division corps level models. I did not include in the chart the

SLIDE 8-9 SUMMARY OF TRENDS IN U.S.



RELACS and TREND models and the DOAE corps and theater campaign models. These are strictly U.S. models on the chart, and some of the names of the players.

Let me show you what has taken place in the battalion and below. Many years ago, probably back in the 1950s or such, we started out with something called Little Monte Carlo Duels, one-on-one tanks. Well, Clint Ancker was very bright. He said, "I can do that analytically" and he created some stochastic duel theory, using simple stochastic processes — one knows how to do that without Monte Carlo sampling. We went from that into very simple Monte Carlo simulations, like CARMONETTE and GLOBAL. Then, roughly in the mid 1960s, we went to high resolution Monte Carlo simulations like improved CARMONETTE, IUA, and then DYNTACS, a very sophisticated, complicated Monte Carlo simulation of battalion level combat. Then we moved into something that I call hybrid/analytic/simulations — some things you do analytically, some things you simulate. You simulate movement, you simulate decision making, but you can analytically handle attrition, acquisition, maybe even terrain line of sight characteristics. So it's a hybrid model between simulatory aspects and pure analytic aspects.

It seems to me that what we are doing, at that level anyway, has gone full cycle. We started out with very simplified analytic structures and have gone to very high resolution simulations, and now we are learning how to come back out of the morass of detail by looking at that detail and coming up with probably in the next two to four years, a perfectly good purely analytic but complex model that will produce similar results and produce simple dynamics. It will therefore be much easier to use without the "10^{30}" sensitivity analysis.

Similar kinds of things have happened at the division corps level. We've run from war games, to computer-assisted war games, to player-assisted simulations, to minimal player participation and fairly detailed simulations like DIVWAG. We have started to work our way out of this now with the FOUR CE model, the DIVOPS model, and are coming to something I call the complex hybrid analytic, the COROPS model. Now my question is, do we know enough about division and corps operations that we can go to pure analytic, or must we really go to high level, high resolution simulation models of division/corps? I don't know the answer to that yet. My guess is we'll probably get some more detailed simulations before we can get back out to the analytic.

Now, let's turn to the topic of this symposium — "the theater level." We have gone from simplified firepower score and other games like JIFFY to firepower score and other simplified analytic models like ATLAS, which is a pure firepower score, fairly like the Lanchester types of model, ATLAS, TCM, and I guess BALFRAM was also set to differential equations.

From there, we have gone to simplified hybrid analytic simulations. That's the whole spectrum, I think, stimulated mainly by WSEG and General Kent, not quite the same four that came out of CAA but LULEJIAN, IDAGAM 1 and VECTOR 1, COMBAT 2 in part — this whole line of nuclear models (Slide 8-9). We have gone from that spectrum, or that generation, the 1972 to 1975 generation, to those that are under development. There's TFECS — that's a model Vector will integrate into CEM-4 for CAA electronic warfare, intelligence, and communications kinds of processes. There's TCOR, I don't know what it is, but I know it exists. It seems to be a move in the direction of more detail. Maybe John Bode can tell us about it later. I have a question mark next to IDAHEX. I don't know its current status, maybe Jerry (Bracken) can comment on it later. At VRI, we've moved into a VECTOR 2, which, as I tried to project, is a very high resolution, simulation type of model with a lot of detail. The next step in that generation is something that Jasper (Welch) has created called CASM. I don't know when that will be done, but as I understand it, it is a much more detailed model than VECTOR 2. I would hope at some point that the trend will be very much as one would see in industrial modeling, for example in inventory theory, in which we went from simple analytic lot-sized formulas to high resolution simulations, and now, we're coming back out of that excess of detail. I would hope we can use the high resolution simulations in theater level warfare to come back out to some simple analytic structures that will provide similar kinds of insights for decision making.

Let me just comment on a few notes here I made during this morning's session. I know Jasper was talking tongue-in-cheek when he said a 10^{30} analysis. The key words he should have put in there were, if we stop to think then we can get away with a lot less than that. I mean, there are lots of sensitivity analyses you can do on the back of an envelope. You don't have to run these big massive loads, when with side analysis, you can see that a lot of things you might investigate don't make any sense. As when, for example, you discover the existence of dominance with certain variables. Anybody who really thinks you want to run a sensitivity analysis is not an analyst. He's a number player and he doesn't belong in the business.

Secondly, I keep hearing everybody say, "Well, we don't like the assumption of that model or the assumption of this model," and then I hear all about the firepower score and the WEI/WUVs and the WUV/WEIs. Well, they have assumptions behind them, too, and I know they're a lot less credible than the assumptions in a lot of the models. I feel that they are just plain idiotic. They may be a nice tool for military people talking about war, as a judgmental device, but they are not models and have terrible assumptions that are a lot worse than the assumptions underlying the models.

I would point out, if Andy (Marshall) would take a closer look, that models have come a long way since 1972. Today, models incorporate a lot of things that he says can't be incorporated. The models don't do the contingency. The analyst thinks about the contingency and puts it in the model to analyze it. I think a lot of the credit goes to Glenn Kent who has stimulated a lot of today's models.

I already mentioned what I thought was a strange inconsistency between people wanting more detail and yet resenting having to spend $120K a year to put three people to the task of understanding the model and how to use it. At the same time, we've got billion dollar investments riding on decisions, and if the models produce a 10% piece of information for inputs in decision making that's good and well worth the money.

I guess what I'm looking for is, the long range perspective in this whole community. I hope we in this group won't look back twenty years from now and say, "Boy, we really had no vision." You know, Frank talked about the phenomenology of what I would call the military science. If you want to understand that it takes some resources, not a great deal, but some money to develop the models, to collect the data. You asked me about the data. Yes, it's going to take three or four years to collect good data — really good data — but it's worth the money. It's not that much money vis-à-vis the money you're spending on the systems you're buying with no information. Okay, we have spent fifteen years in the Army collecting data like hit probabilities and kill probabilities. It took time to do that, but we ought to collect it for the theater models.

I don't see any money being spent on verifications. You say the models are no good but that's an experimental question. That's not a question based on judgment. It's a question of data, and there is data. WSEG has had it for years. British was data. I have submitted five proposals to this community to test any of the models with it and nobody wants to test the models, but they're willing to stand up here and say they're no good, they don't predict, they don't do annything. How do you know? If there is data, and you're not willing to spend a dime to see if they do predict anything, how do you know? I just think we need some more long range perspective in this community, and if we don't get it it's going to die. Maybe that's okay with a lot of the users, but I think they ought to understand that if they want some benefit from it they have to put some money into it.

Dr. Bracken: I might add a comment to what Seth (Bonder) has been talking about, about the status of model development. At IDA, our model development efforts have shrunk to near zero. Paul Olsen, who hadn't intended to come to this conference, but who is here, is a developer of IDAHEX and he spends some of his time on IDAHEX development but most of it on IDAHEX application.

Ed Kerlin will talk about TAC WAR later in this session. His group has been concentrating mostly on data base development. The only major initiative I know of is CASM, which has not been too active recently, to the best of my knowledge. So I agree with Seth that it's time for a reexamination. John Bode commented at lunch that it was time to build data bases, and perhaps it is.

We've been doing studies and analyses but we're now in very much a hiatus as far as model development in the United States is concerned. I asked Reiner Huber and Klaus Niemeyer what they were doing in Germany. They have about twenty to forty people involved in model development in IABG.

The next speaker is Dr. Alan Karr. Alan Karr was with the Institute for Defense Analyses for a couple of years, and he's been a professor at Johns Hopkins for about three years. He has been systematically looking at four models, IDAGAM 1, LULEJIAN 1, VECTOR 1, and the CEM, and has done a critique of them. After Alan Karr speaks, Lieutenant Colonel Lanny Walter will speak. He and a team from WSEG have also done a critique of the four models. So this is an attempt at a systematic look at the four models developed at roughly the same time. Now, let me turn over the microphone to Alan Karr.

9 — *A Critique of Four Theater-Level Models*

PROF. ALAN KARR
Johns Hopkins University

Four theater models (CEM, Lulejian, IDAGAM, Vector) and their underlying assumptions for representations of attrition and FEBA movement.

Prof. Karr: In the interests of terminating this session at a reasonable time, I will try to cut out some of the things that I had to say. Seth Bonder has certainly helped me in that by preempting one-fourth of my talk, and Jerry Bracken with a small fraction. So I'll consider those discussions part of this presentation.

What I will describe very briefly here has been an effort funded by IDA to study in some detail, and perhaps from a different point of view than some others had tried, the assumptions underlying various theater level models, specifically the CEM, the LULEJIAN, IDAGAM, and the VECTOR series. The VECTOR 1 critique has been out for some time, and I'm currently working on a critique of VECTOR 2.

There are several reasons for doing these critiques. One is to produce a different exposition, a different viewpoint of these models than those that are incorporated in either the model documentations themselves, or in the Four-Model Comparison Study. In fact, when you hear about the Four Model Comparison Study, you may have a hard time believing that it discusses the same things that I'm talking about. I think this is useful. All of us realize that one exposition of some material may be utterly incomprehensible, while another exposition may be relatively easy to follow. So, one of the goals in this work has been to produce useful syntheses and descriptions of these models so that they will become accessible to people who, for one reason or another, have difficulty with the model documentation.

The reason I have emphasized the role of underlying assumptions, in the models, and specifically the role of assumptions underlying attrition equations, is that it gives me an understanding of exactly what the limitations in the models are so that I can try to help others who read these critiques to understand what those limitations are. After all, assumptions are limitations, and if one is to use the model intelligently then those limitations really ought to be as explicit as possible and as well understood as possible. The assumptions can't really tell you exactly when you can use the model, but they certainly can give you information about situations with which it's totally incompatible and that, of course, is better than nothing.

In addition, although Seth (Bonder) has mentioned a few things about availability of data, I think some of the other methods that one might think of for comparing models are perhaps not easy to use in this situation, such being empirical methods. Any sort of experimentation seems to be out of the question, so one basis on which models can be compared is by looking at the assumptions that underlie them. To do that, of course, you have to know exactly what the assumptions are. Now, it's quite true that there are certain mathematical structures that may be led to by one or several sets of assumptions, and one hopes there aren't too many structures around that are led to by nil sets of assumptions, although I think there are a few of those too. Nonetheless, these critiques have been at least an effort to say something very explicit about the assumptions.

In general, in these critiques I have criticized things on a more or less absolute basis in the sense that there is not too much model-by-model comparison. There is a great deal of that, of course, in the Four-Model Comparison Study about which you'll be hearing. So perhaps one ought to temper a lot of the things that I've said with the realization that they are absolute criticisms, and I am well aware, although perhaps not to the extent that some of the users or potential users of models are, of the practical considerations involved.

One other thing I've attempted to do, where I felt I was able to, was to suggest alternatives to some of the structures that were used. Now, what I'll do with the remaining time is attempt to give you more of the flavor of the evaluation and critique that I've done for each model than to be complete at all.
Specifically, I have said something about what seem, on the basis of my reading of the documentation, to be the stated purposes of the model. This is not particularly evaluative. I have attempted to portray accurately what I think the model documentation was saying that the model is intended for. I have then stated what I think seem to be the emphases of the model and then just as sort of a sample of some of the things that have been done, some statements concerning attrition and FEBA movements. So, I can quickly go through these models in alphabetical order.

Slide 9-1 — CONAF EVALUATION MODEL (CEM)

STATED PURPOSES:
- Comparison of different force structures
- Analysis of combat decisions on missions of subunits and allocation of supporting resources

EMPHASES:
- Complex, hierarchical decision-making structure
- Renewable resources
- Detailed resource accounting

ATTRITION:
- Based entirely on kill potentials and exponential equations

FEBA MOVEMENT:
- Computed from force ratios

The CEM (Slide 9-1) states, as do most of these models in one form or another, that it's intended to be used for comparing different force structures. Explicitly, it does state that it might be useful for analyzing effects of decisions. To that end, it has within it, and, in fact, one might say that it consists almost entirely of, a rather complex hierarchical decision making structure in which decisions at higher levels constrain alternatives of units at lower levels. There is also significant emphasis, in my opinion, on renewable resources and on the resource accounting.

The attrition in it is based almost entirely on exponential-looking equations which have in them kill probabilities — kill potentials. FEBA movement is computed from force ratios.

The IDAGAM (Slide 9-2) documentation is not so specific about exactly what the model is supposed to do except relative to other models. If you read the report, it gives one the impression that the model was created as a result of some sort of long-term discontent with previous models, which I think is probably true. One of the things that I think the developers of the model attempted to do was to put the air and ground combat levels at comparable levels of detail. Toward that end, they have emphasized the air combat model. It seems to me to be somewhat more detailed, if not significantly more detailed, than those in at least the LULEJIAN and VECTOR 1 and CONAF models. They have very carefully, I think, integrated the air and ground combat but on a simplistic level, not on the same level of detail as VECTOR 2.

One of the emphases is on a number of different attrition equations. At ceratin points, the user is given a choice in the model of what equation he wants to use for computing certain attritions. Like a lot of other potential strengths of various models, this is a strength and at the same time possibly a significant weakness. There's lots of potential for abuse here is you don't know what you're doing. The same applies to the possibility of the alternative force ratios.

So far as attrition is concerned, the air combat attrition is where the choice exists. You can pick anything from Lanchester square to Lanchester linear — four different types of exponential equations. For ground combat the attrition is a Lanchester square attrition except where it is scaled on the basis of casualties. Casualties are computed in absolute terms from force ratios. People movement is also computed from force

ratios but it allows the representation of differing concentrations of attacking air forces depending on the relative air and ground advantages of the attacking side.

Slide 9-2 — IDAGAM I MODEL

STATED PURPOSES:
- Improvement in general over previous models
- Comparable levels of detail in both ground and air combat

EMPHASES:
- Air combat model and integration of air and ground combat
- Alternative attrition equations
- Alternative force ratios

ATTRITION:
- For air combat, choice among alternatives of exponential and Lanchester form
- For ground combat, scaled Lanchester Square with casualties computed from force ratios

FEBA MOVEMENT:
- Computed from force ratios; differing concentrations of attacking air forces permitted

Slide 9-3 — LULEJIAN-I MODEL

STATED PURPOSES:
- Generate approximately optimal campaign-long allocations of major resources
- Assessment of net combat capability
- Evaluation of alternative force structures

EMPHASES:
- Optimization structure
- Effects of supply movements and shortages
- Joint, iterative calculation of ground attrition and FEBA movement

ATTRITION:
- For air combat, exponential equations
- For ground combat, exponential equation with separation distance in denominator of exponent

FEBA MOVEMENT:
- Computed with attrition so that all resources on each side incur same ratio of actual to acceptable losses

One of the principal purposes, and therefore one of the emphases, of the LULEJIAN model seems (Slide 9-3) to be to produce some sort of otpimization in allocation of campaign resources. It has a complex sequential game structure built onto it, but one could perhaps choose to ignore the structure, or make it relatively less important. It would be useful with that superstructure even for assessment of combat capabilities and alternative force structures. The emphasis, as I noted before, is the optimization structure.

Second, is that there is a rather detailed representation of supply movements and shortages. There is a more or less explicit, although not geographically explicit network of supply conduits, which can be attrited and built up, and so forth.

Another emphasis is a joint methodology which at the same time calculates both ground attrition to principal weapon systems and FEBA movement. The air combat is handled by exponential-looking equations. The FEBA movement and ground combat are calculated together by choosing separation distances which, as far as I can tell after reading through everything, do nothing but insure that on each side all of the resources suffer the same ratio of acceptable-to-actual losses, or, actual-to-acceptable. Furthermore, the ratios on the two sides are related by being one — I guess inverses of each other.

Slide 9-4 — VECTOR-2 MODEL

STATED PURPOSES:

- Provide information useful in making net assessments and general purpose force tradeoff analyses, and in studies of strategy and tactics

EMPHASES:

- Coordinate locations of maneuver units and aircraft
- Attrition computed from physically definable inputs
- Tactical decision rules
- Level of detail and completeness

ATTRITION:

- For air combat, equations of multiple engagement binomial form
- For close ground combat, discrete approximations of differential equations that explicitly represent lines of sight, range, target acquisition and selection, and time-to-kill

FEBA MOVEMENT:

- Equal by definition to movement of front line maneuver units

Finally, I have not said anything about VECTOR 1 since I'm now in the process of working on VECTOR 2 (Slide 9-4). It seems superfluous at this point to say very much about VECTOR 1. I think, Seth (Bonder) has given a rather better description than I could give, especially in a short time. In fact, I am somewhat pleased that there is at least some resemblance between what I have said and what Seth is saying.

I think that concludes what I have to say. I really don't feel I can go into too much detail here. The study had no formal name, but the four critiques were issued as IDA papers, and I believe they are available at DDC.

Dr. Bracken: Now, I'd like to introduce Lieutenant Colonel Lanny Walker. Lanny is now stationed at Fort Dix. He was an analyst with the Weapon System Evaluation Group (WSEG) for several years and led the team that did the Four-Model Comparison Study.

10 — Four Model Comparison Study

LTC LANNY WALKER, USA
(formerly of WSEG)

Ranking of CEM, Lulejian, IDAGAM and Vector in their modeling treatment of military functional areas

Col. Walker: As Alan Karr has already stated, our model comparison study involved four theater level models:

CEM-IV
IDAGAM-I
LULEJIAN-I
VECTOR-1

All four of these are deterministic models. They are related primarily in that they are of approximately the same generation. Three of them were developed in conjunction with, or sponsored by, WSEG except for certain portions of IDAGAM, and CEM was developed independently for CAA of the Army by GRC.

We jumped aboard a moving train in this study. As you have undoubtedly gathered by now, there have been, and continue to be, major modifications to VECTOR 1. Also, LULEJIAN 1, as it existed at the time of our evaluation, still exists, but there have been major modifications of various sorts in LULEJIAN 1 for specific applications.

At any rate, we attempted to look at these four models with an idea of trying to give potential users of theater level models, and of one of these four in particular, some better basis than just a broad reading of the documentation of the models on which to judge the selection of a model for a particular purpose. The purposes that we have studied are simple:

- To assist potential model users in understanding the capabilities and limitations of the four models.
- To provide a guide for the selection of models for specific applications

You'll notice we made no attempt to determine in any way whatsoever the best model of these four. We wanted to provide assistance to the user community in selecting a model, and, more generally, in understanding the four models.

We used certain ground rules. One of these was that it was a three-agency evaluation, as it were. I was with WSEG and was the project officer. There was a representative from SAGA and a representative from PA&E. The three of us actually made the evaluation. We based the evaluation on the documentation, as did Alan Karr in his study. We made no attempt to get the models up on any particular computer, or to run comparative runs, or, for that matter, to do any sensitivity analysis on the models. If previous runs were available, we tried to make as much use of those as practical, but most of the analysis that we did is based on the documentation as provided for each of the models.

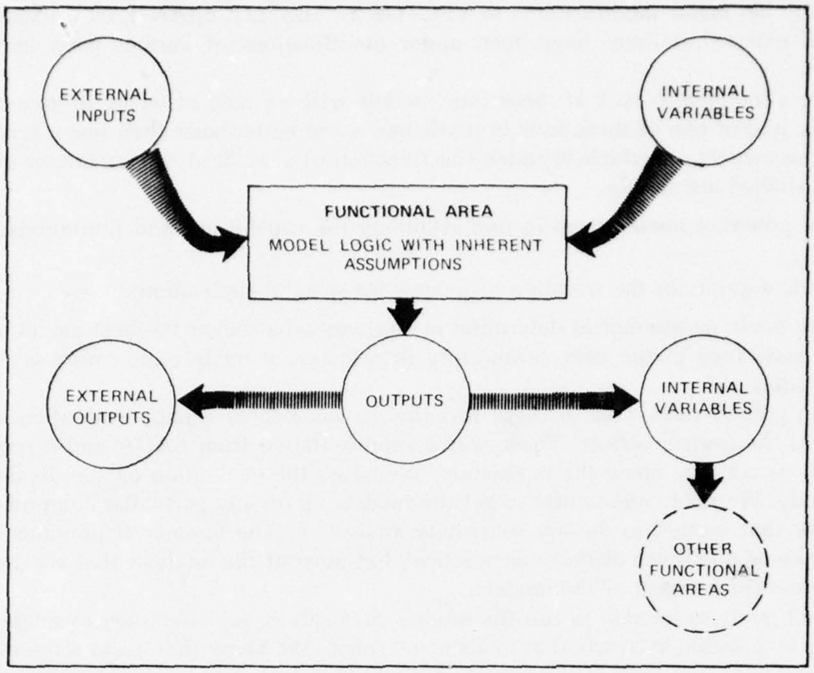
Since we weren't going to be able to run the models themselves, we attempted to relate these models as honestly and clearly as possible in a way that made some sense. We knew that, from a theoretical standpoint, we couldn't just stand back and look at the model as a whole. So what we came up with were a number of functional areas which really have more to do with what we perceive as taking place in theater combat than with the particular structure of the individual models. So, what we did was break each of the models down into these functional areas. In many cases they pretty much corresponded to the way the model was structured itself, but in some cases they didn't. We tried to look at each of these functional areas as a separate model element and to compare among the models each element in turn, each functional area. As I said, we did not make an evaluation of the models as a whole, nor did we make a specific evaluation of the way that each model put together the various functional areas that are represented. We made an implicit kind of evaluation of that kind, and I'll speak about that in a minute.

Slide 10-1 — MAJOR FUNCTIONAL AREAS

Organization and Missions of Forces
Geographical Theater Structure and Forces
Force Interactions
Ground-Ground
Ground-Air
Air-Air
CAS
Interdiction
Logistics
Unit Replacements and Individual Replacements
Intelligence

The way we went about looking at the functional areas is depicted on the chart (Slide 10-1). We looked at the sets of external inputs. That is, inputs to that functional area which are provided from outside the model — data, if you will (Slide 10-2). Now, as was mentioned earlier, that does not mean necessarily that these external inputs are pure in any sense. Quite likely they are the output of some other model. These inputs may be simple or they may be fairly complex, but at least they are external with respect to the model that we're looking at. And, as an input, they are, of course external, to the functional area.

SLIDE 10-2 EVALUATION METHODOLOGY



Then, we looked at the internal variables with an eye toward answers to the questions "do they make sense? And, "Are these the things that we believe actually affect the functional area in real theater combat?" We looked at the way the functional area put together these inputs, the assumptions involved, and the model logic, again, to see in our judgment whether they made sense.

Then we looked at the other end, the outputs, the external and internal variables that came out of a given functional area and then were passed to other functional areas.

Of course, in looking at internal variables coming into the model and internal variables that were produced in any given functional area, we were implicitly looking a little bit at the structure of the model as a whole and the way the various functional areas were interrelated.

Up to this point insofar as we tried to determine from the documentation precisely what assumptions were built into each portion of each model, we were able to keep it fairly objective. Unavoidably, there were some subjective judgments but we tried to keep treatment relatively free of them. We felt that, once we had described as well as we could each of the functional areas we should make some attempt to evaluate how well each model did what it purported to do in each of the functional areas.

These are the essential criteria that we attempted to use, all of them judgmental.

- Degree of model aggregation and abstraction.
- Degree to which significant activities or functional areas are represented.
- Degree to which a variety of operational concepts can be reflected in the model decision processes.
- Degree to which model decisions are influenced by important aspects of the ongoing combat situation.
- Degree to which individual types of units, weapon systems, and other elements can be represented discreetly.
- Degree to which weather, terrain, and other environmental factors can be represented.
- Degree to which results are obtained from internal calculations, rather than look-up tables.
- Degree to which the mathematical expressions accurately describe the physical process being represented.
- Value of unique features, e.g., VECTOR input tactical rules, LULEJIAN "optimization" routines.

Some of them are easier to get at than others.

Slide 10-3 — MODEL FUNCTIONAL AREAS — SUMMARY OF RANKINGS
(Scale 1-10)

Functional Area	CEM	IDAGAM	LULEJIAN	VECTOR
Organization and Mission	1	5	4	3[1]
Geographical-LOC	1	4	4	3
Interactions				
Ground-Ground	5	2	4	2[2]
Air-Air	7	2	3	4
CAS	6	3	2	2
Interdiction	7	5	4	5
Ground-Air	8	3	4	4
Logistics	7	7	5	4[1]
Unit/Personnel Replacements	4	7	5	5[1]
Intelligence	8	10	10	8[1]

[1]Assumes that appropriate tactical decisions rules are programmed by user.
[2]Refers only to the structure used for an attack on a hasty defense.

In our report, which was published as a WSEG report, we summarized the strengths and weaknesses as we perceived them for each model in each functional area. These were listed along with a rating, a value judgment, on the part of the members of the team as to how well each model performed that function in relation to the other models. We devised a very simple, arbitrary rating scheme from one to ten, the higher the number the less we thought of the way that that particular model did that particular thing (Slide 10-3). The numbers had no meaning except in that context — they couldn't be added — and we really didn't do too well, I think, in relating the numbers from one functional area to the other. However, I think we did do reasonably well in relating the models one to the other in a given functional area.

Slide 10-4 — GROUND-TO-GROUND INTERACTIONS
CEM (5)

STRENGTHS:

- Engagements and attrition calculations take place at the brigade (or its equivalent) level, or lower.
- Attrition of 15 types of tanks and 12 types of APC's is represented individually.

LIMITATIONS:

- Sums of directed firepower scores are used in attrition calculations.
- An input K-factor is used in the exponential attrition equations to relate firepower scores, numbers of targets, and target attrition.
- Losses cannot be attributed to the weapon type which inflicted them.
- Attrition of non-armored weapons (e.g., anti-tank weapons) in maneuver units is not directly computed. It is assumed to be proportional to dismounted personnel attrition.

Let me show some examples of the level at which we listed strengths and weaknesses, or strengths and limitations, and as some of you may note here in slides 10-4 through 10-7. We've called them strengths and limitations on these slides, but in the report we called them characteristics to remove some of the element of value judgment that you see here. The report itself consisted of a portion with characteristics of each of the models at about this level of detail, and then another larger section which went into detail on the inputs, assumptions, outputs, and descriptions of how the model works, and, lastly, a summary sheet in which we say the model should not be used except in relation to detailed reading, which describes each of these functional areas. We tried to use criteria that I showed earlier in arriving at these general ratings. As I mentioned before, I think they have more validity horizontally, than vertically, but we tried to get as much validity vertically as we could. The report was published early this year and is available at DDC.

Slide 10-5 — GROUND-TO-GROUND INTERACTIONS
IDAGAM (2)

STRENGTHS:

- Attrition equations are applied to individual firing and target weapons, by type to obtain "potential" losses. Attrition can be attributed to the system inflicting it.
- A variety of optional techniques are available to scale "potential" attrition to obtain "actual" attrition. Overall "actual" attrition can thus be made a function of force ratio.
- Various optional methods for computing force ratios are available.
- The Lanchester forms of the attrition equations are well-known and transparent. Attrition coefficients are modified each day to account for changes in allocations of fire.

WEAKNESSES:

- Attrition calculations are made at the sector (corps) level. Therefore:
  - Forces are assumed to be homogeneous across an entire corps front.
  - Maneuver and mobility below corps level are not modeled. Their effects must be determined outside the model and reflected indirectly in appropriate model inputs, e.g., kill rates, scaling factors.
- Standard allocations of fires must be input. Only linear variations from the standard are computed internally, based on target inventories.

Slide 10-6 — GROUND-TO-GROUND INTERACTIONS
LULEJIAN (4)

STRENGTHS:

- A direct relationship exists between FEBA movement, the level of attrition, and the number of battalions committed to combat. Attrition caused by air and ground supporting fires is considered along with attrition caused by maneuver unit weapons.
- The ability to locate opposing force elements is modeled. As the locations of opposing elements become more accurate, the effectiveness of weapons increases.
- The effectiveness of maneuver unit weapons can increase as the range to the target decreases.
- The overall results of ground interactions, reflecting supporting CAS, are used as the payoff criteria for choosing aircraft allocations.

WEAKNESSES:

- Weapons are aggregated for attrition calculations. For example, the total anti-tank potential of ground weapons and aircraft is applied as a single value against a single, average generic tank. Results are subsequently disaggregated, so that the number of weapons of a specific type which are destroyed is proportional to the number which are vulnerable.
- Attrition of only three primary elements of maneuver units are calculated directly: tanks, APC's, personnel. Losses of other weapons (e.g., anti-tank weapons) are assumed to be proportional to one of the three primary elements.
- Allocations of fire to target weapons are implicit in the input potential capabilities of firing weapons.

Slide 10-7 — GROUND-TO-GROUND INTERACTIONS
VECTOR (2)

STRENGTHS:
- Combat is modeled dynamically using small (e.g., 1 minute) time steps.
- Interactions among weapons, by type, are modeled in detail, using basic weapons performance data. Such processes as movement, acquisition, target selection, target engagement, and target destruction are represented explicitly.
- Allocations of fire are calculated internally, on the basis of input priorities.
- Losses can be attributed to the weapon system which inflicted them.
- Supporting fires (e.g., artillery, CAS) are dynamically represented, and integrated into the ground combat model.

WEAKNESSES:
- The dynamic ground combat model is used to represent only one type of activity for each side: an attack on a hasty defense. The results of other ground combat activities are obtained from look-up tables.
- The relationship between the results of separate, battalion-level engagements and results at the theater level must be defined and programmed by the user. (From the standpoint of flexibility, this could be considered a positive feature.)
- At the battalion level, a relatively small number of tactical scenarios is available.
- There is no representation of adjacency of on-line battalions. Locations are in terms of sectors only.

Okay, that completes my presentation. I do have a little bit more information about some of the other functional areas, at about the level that we had for ground-to-ground interactions in case anyone is interested.

Dr. Bracken: Now, I'd like to introduce Ed Kerlin of IDA. Ed was the project leader of the ATLAS, a model development at Research Analysis Corporation in the early 1960s. He was a member of the IDAGAM development group at IDA, and for the last few years he's been the project leader of a study called TAC WAR, which is modeling conventional tactical nuclear and chemical warfare at the theater level. For the last nine months he and his colleagues have been building a data base for the TAC WAR model. That is the study that he will talk about today.

11 — Theater-Level Modeling of Conventional, Nuclear and Chemical Warfare

DR. EDWARD KERLIN
Institute for Defense Analyses

Description of TACWAR — a theater model of conventional, chemical, and tactical nuclear warfare.

Dr. Kerlin: The work I want to talk about today is what we've been doing for JCS/SAGA, and most recently J-5, Plans and Policy Directorate.

I think that all the model builders are aware that the community has had a number of conventional war models available to them at the theater level (Slide 11-1). If you look at them, as we've already seen by various reviews today, they occur at different levels of resolution, they have varying methods of assessment, and, of course, they're widely different running times when you consider the kind of data that is in them. Our interest was to try to produce for the JCS a theater level model that played the conventional war but one in which you could incor-

SLIDE 11-1 BACKGROUND TO THE STUDY

- ***JCS/SAGA (AND OTHERS) HAD AVAILABLE -***
  1. **CONVENTIONAL WAR MODELS**
     - **AT DIFFERENT LEVELS OF RESOLUTION**
     - **WITH VARYING METHODS OF ASSESSMENT**
     - **WIDELY DIFFERENT RUNNING TIMES**
  2. **NUCLEAR ASSESSMENT MODELS**
     - **MOSTLY HIGH RESOLUTION MODELS**
     - **MONTE CARLO ASSESSMENTS**
     - **COMPUTER ASSISTED OR LONG RUNNING**
     - **LIMITED SCOPE OF TARGETS**
- ***THE JCS AIR BATTLE STUDY PROVIDED THE IMPETUS FOR***
  1. **AN INTEGRATED CONVENTIONAL-NUCLEAR MODEL**
  2. **DYNAMIC ASSESSMENTS AT THE THEATER-LEVEL**
  3. **EXPECTED VALUE (OR DETERMINISTIC) ASSESSMENTS**
  4. **NUCLEAR EMPLOYMENT OPTIONS**
- ***OTHER INTERESTED USERS INDICATED***

  "THE LOW RESOLUTION MODEL CONCEPT, INTEGRATING CONVENTIONAL AND NUCLEAR WEAPONS, BOTH AIR AND GROUND, ASSESSED AT THE THEATER LEVEL, WAS OF MAJOR INTEREST."

porate the concepts of a tactical nuclear environment. So then we went back and looked at the nuclear assessment models, and found that they were mostly very high resolution models. For example, the SATAN model, which has been used in JCS for a number of years, is very detailed. It plays platoon-sized units on the battlefield, and these become the targets for the nuclear weapons. SATAN is a Monte Carlo assessment, and it takes a lot of running time because you have to do iterations.

Also, as Seth (Bonder) commented, there are models like DIVWAG that have nuclear components in them, but these are either computer systems or war games and they have very long running times as well.

Furthermore, the nuclear models we looked at were very limited in the scope of targets that they considered. There was not a model around that would, say, start from battlefield targets in the forward area, go deeper to reinforcement second echelon targets, go deeper to supply nodes, go deeper to airfields, and that would run the gamut of the entire targeting concept at a theater level.

About the time we were starting to build a theater level tactical nuclear war model, the JCS was asked to do an air battle study. There, the idea was to look at the contribution that tactical air would make in a tactical nuclear environment in Europe. How do you measure that? How do you analyze that particular situation? Certainly, it may or may not start as a nuclear war. It may start conventionally, and escalate in some sense to a nuclear environment. They found very quickly that an integrated conventional/nuclear model would have been very helpful to have at that time to do the study. Lacking that, they sort of lashed together the existing models they had onboard — the IDAGAM 1 theater level conventional model and the SATAN model I've already alluded to.

The one problem with these models is that it takes a long time to do one iteration, something of the order of an hour and a half to two hours. If you do statistical validity, you're talking of the order of maybe 20 hours of computer time, which is a lot of time. So our view of the problem was that we should somehow be able to reduce this running time and yet get the spectrum of targets that we wanted into the model. That either meant cut down the amount of detail in the model or go to an expected value or deterministic assessment.

The other area that was of concern at this time was the recognition that most nuclear models were driven by user input — the user picked out the target, or the target was picked out for him by some target acquisition model, then he directed a certain weapon on that target. Nuclear models are not easy to play, particularly in terms of the concept of what the Army has called "escalation boundaries." How can you keep nuclear warfare limited? Can you look at restricted use of nuclear weapons? Can you look at more extended use of nuclear weapons? Just how do you play the employment of nuclear weapons in a theater environment? Those four items were, in a sense, what the air battle study opened up to us — they became the initial basis for our development of the theater level tactical nuclear war model.

Following our initial efforts in that area, the J-5 became very interested in knowing where U.S. NATO forces would be in a defensive situation if the enemy should suddenly begin to use chemical weapons. We don't propose to use them offensively, of course, but the question was asked, "where do we stand if we're hit with them in a defensive role?" So, we were asked to try to incorporate that modeling concept into this theater level model as well, enlarging to the arena of chemical munitions.

I'd like to make just one point regarding the lower portion of this slide. Initially, we were not sure that the low resolution theater level model was the way to go. But we had talked to a number of people who had been building models, and, over the years, had attended symposiums like this one, and the consensus was that a low resolution model that integrated these conventional and nuclear munitions delivered by air and ground throughout the theater was really needed. There just wasn't a model around that could do this. We realized the difficulties we were making for ourselves. We have never fought a nuclear war, so how do you get all of the parameters? How do you get the assessment features into a model of this nature? This was perhaps our most difficult problem.

To review just briefly: we'd already seen some of these models before. We spent considerable time going back over the literature, the documentation, talking to people who used these models, and became very familiar with what they did in terms of assessment and running times. Shown below is a list of theater-level nuclear and nonnuclear models:

Nonnuclear Models (Theater-Level Combat)

- QUICK-GAME → ATLAS
- TOTEM → TALLY/TOTEM
- TCM → CEM → ATENA
- GACAM → IDAGAM I → IDAGAM II
- FRAM → BALFRAM
- BATTLE → VECTOR I → VECTOR II
- LULEJIAN I → LULEJIAN II

Nuclear Models

- NAR/TAR I, II, III → TANREM
- SATAN I → SATAN II → SATAN III
- DWEEPS
- TANDEM
- UNICORN*

Combined Nuclear/Non-Nuclear Models

- COMBAT II
- TACWAR*
- (MINTSIM)

*Includes chemical warfare assessments.

I've tried to show them in time frame of development down the column and what they evolved into (arrows). Then, our assessment of the nuclear damage assessment models is shown — I'll call them that instead of really dynamic models. Finally, at the very botton, I show two other studies that were in process as we were doing our study — the BDM model COMBAT 2 and the Army concept that General Research Corporation was doing, the MINTSIM concept. That is shown in parenthesis to indicate that it was a modeling specification and not a true model.

Another basic point was that we did not want to try to re-invent the wheel in terms of modeling, so we looked carefully at the models that I've listed. A reason for selecting those models is that we're more familiar with them than with some of the others we could get from documentation plus they seemed to provide the kind of information and structure that we felt would be needed in our models.

I show the tandem model because it was very useful to us in two ways: it gave us a very detailed data base on airfield locations and the structure of airfields on both sides of the NATO Warsaw Pact area, and it provided us with a population data base. We modified that set of data somewhat to accommodate our model, but it was very helpful to us.

On another point, Seth (Bonder) has already talked about going into detailed simulation, getting immersed in detail, but then hopefully coming out with analytical expressions to reduce analysis time, running time, and data requirements. We've tried to do that in both the chemical and nuclear assessments.

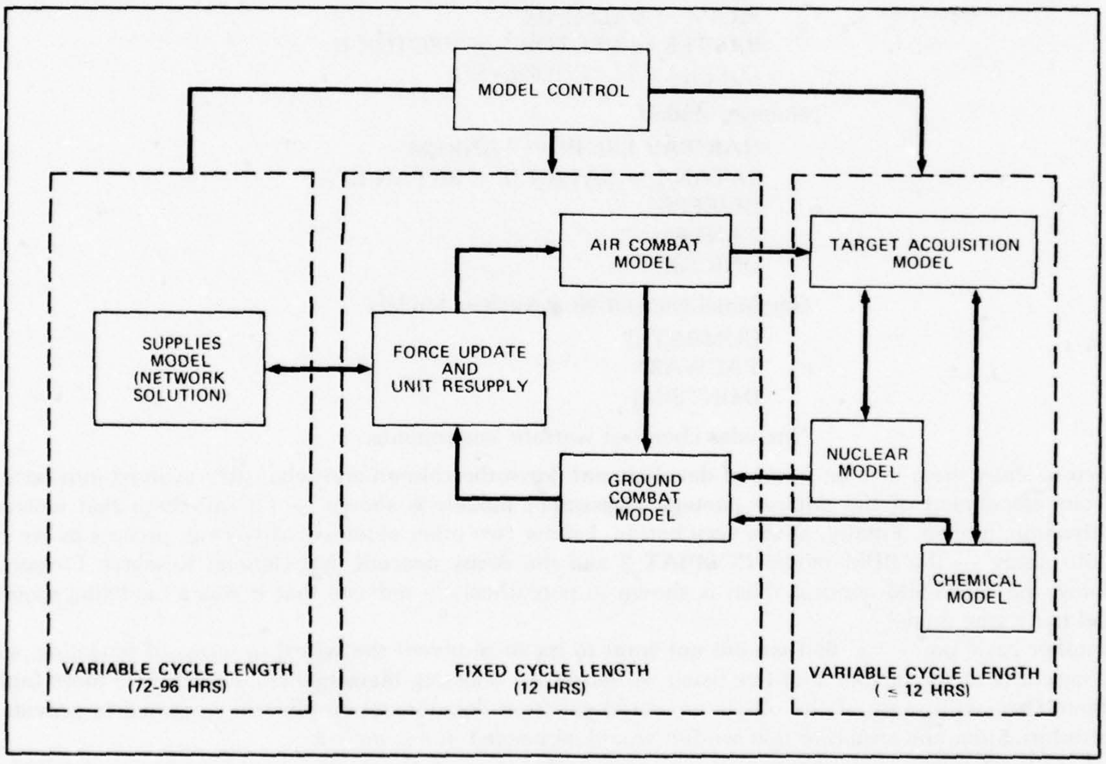
If you go back to the models that I've talked about here — SATAN 1, SATAN 2, SATAN 3 — the detailed assessment methodology in those was just too detailed for us and too long running. What we tried to

do for the nuclear case was to look at the methodology they used and to come up with analytic functions that would give us close approximations yet cut down on the tremendously long running times.

For the chemical assessment, we spent some time with the folks at Aberdeen. They have access to a number of very long running models that can generate the distribution and the diffusion of a cloud downwind as it explodes at a certain altitude and then as the winds carry it down. How does this cloud develop and what's the concentration of that cloud at various points downwind? If you get involved in a model that considers all the parameters related to these questions you are running into hours and hours of computer time. Well, we took the chemical assessment methodology Aberdeen had and did the same thing with it as we did with the nuclear case. We ran the model, got the outputs, and then used our analytic functions to approximate the model results to get the same kinds of dosages as a function downwind, given the parameters of inputs like height of burst, the kind of agent, droplet size of agent. I'll try to cover those a bit more in detail later.

Anyway, this was the basis for our modeling concepts. We wanted the model to exist in one of two or three ways, actually. It wasn't always necessary to run this model as a conventional/nuclear/chemical model. Since it was conceivable that one might want to try to compare its results to other theater level models, we have a mode that we call the "conventional mode" (Slide 11-2) It plays the entire theater level conventional air war, with the features that Lanny Walker has talked about — the air-to-air, the air-to-ground, the ground-to-air, and the close air support interdiction of air base attack.

SLIDE 11-2 OPERATIONAL STRUCTURE OF THE TACTICAL WARFARE MODEL



The interaction of air combat with the ground combat model then is depicted on NATO corps sectors, using combat units of division size. In our sense, they're division-sized because we are core limited. If the user had more core space available, the division-sized combat units could possibly be reduced to brigade size, but I don't think it would be possible to go down to a level as fine as battalion size. Then, in this model the divisions fight on a fixed cycle length of 12 combat hours. At the end of 12 hours there is either a force update of replacements — people and equipment — or resupply to that unit. Thus you can run this model basically on a 12-hour cycle, portraying only a conventional war.

We feel that the core of the air-ground attrition is very similar — in fact, exactly the same as the IDAGAM 1 attrition. However, in our portrayal of ground combat, we tried to fix specific unit identities because nuclear strikes on a specific division are going to render that unit ineffective, and we had to know which one it was going to be.

Now, the level of resolution on nuclear strikes is much finer than that. We subdivided the division into basically company size, or battery size units. The user then has to structure his forces into maneuver units of company size and artillery units of battery size. We've allowed ourselves seven subunit types to structure that division. Now, you may say that doesn't sound like very many, but most of the models that go into that level play even fewer. Again, if you have ample computer core size you may want to play more. We were limited in what we could do, and I'm not sure that you really need more than what we had. I think you can portray all the firepower capability of a unit as well as its maneuver capability in artillery units, in air defense units, in headquarters units, possibly of two types and even support units. So that is how we've structured each division in terms of its subunits.

One other component that was of interest to us was supplies. There is always great concern in theater level models about what to do about supplies. How detailed do you play them, how grossly do you play them, and do you play them at all? When we first started working on the ATLAS concept, we gave some semblance of logistics resupply to users. I would say 80% of the time the user ignored the ATLAS resupply-capabilities and gave the theater as many supplies as it needed to operate without being ineffective or bothered by supplies. Well, we tried again, this time by looking at the supply network in the theater as a series of nodes and arcs over which supplies had to move because some of these supply nodes could be hit with a nuclear strike and be rendered ineffective, or wiped out. So damage to the supply network should have a more significant impact on the battle under those circumstances than in conventional warfare. In this structure, we incorporated a supply network and we tested the operation of that network based on the demands of the various nodes. This was done by the very rapid running Trans-shipment Algorithm that we borrowed from the National Bureau of Standards (NBS). This algorithm operates on a different time period than the ground combat — something like three to four days for this kind of theater resupply. However, it can be run in less time if you want.

The next feature we added to the model brings in the two model concepts I talked about: the nuclear model and the chemical model. Slide 11-2 shows a target acquisition model as well. In the conventional mode, which I spoke of, we do not play detailed target acquisition to generate targets for these conventional weapons. That certainly is a limitation when you try to hit individual units with chemical or nuclear weapons. So we felt that there had to be some representation of target acquisition.

The way target acquisition is played in most models is through a probability detection matrix that says at a certain distance from FEBA I'll detect certain types of units with a given probability. Now that set of probability values has been developed over the years, and while it may be valid, there still remains a problem. Suppose you reduced the capability of your force tremendously? How do you modify those probabilities of detection? They obviously don't stay constant. If you bring more force into the theater, how does that affect detection probabilities? How do they go up? There is really no way of going back and saying how do we modify that. So, we built a model that represents all of the sensor systems, ground and air, tactical air, Army air,— at least we have the structure. I'm not so sure, however, that we're able to get the data to support that structure to the level that we would like. Knowing that, we made an option in the model, and I guess this is a fall-back position. We allowed the user to use the probability detection matrix, so, at least, he can make a comparison between that and the other models that are not of this nature.

Well, instead of talking about all the models individually, let me speak only to the chemical portion — the resources that we consider in the model, and therefore the targets for all of our weapons. For our division sized units, we portray individual units and the artillery rockets and missiles that are nuclear-capable, chemical-capable, or conventional-capable, or they can have duel capability. I've talked about subunits in which the divisions can be restructured and have shown that each division has the following resources of people, weapons, and supplies (Slide 11-3).

Tactical air resources consider aircraft of a number of different types. We are currently playing seven types of aircraft. We could play more if we had the space, but we do not have the space. As you can see there are other resources that we play in the model.

I'd like to move on to the chemical model process (Slide 11-4). The important point about the chemical model is that there are three aspects to what is happening: you want to consider a certain employment doctrine of these weapons, you want to be able to use that doctrine to assign weapons to targets, and then, obviously, you have to have a way of assessing the effects of those weapons on the targets that they've been

Slide 11-3 — RESOURCES CONSIDERED WITHIN THE TACWAR MODEL

A. GROUND COMBAT RESOURCES
1. Division (or Brigade) Size Units
2. Nuclear and Chemical Systems: Artillery, Rockets, Missiles
3. Division Subunits: Company, Battery, HQ, etc.
4. People, Weapons, and Supplies for Each Unit
5. Surveillance Sensors

B. AIR RESOURCES
1. Tactical Aircraft with Type Munitions
2. Reconnaissance Aircraft with Sensors
3. Aircraft Shelters at Actual or Notional Airbases
4. AAA and SAMs as Point or Area Defenses
5. Supplies at Airbases

C. OTHER RESOURCES
1. Supplies at Nodes of resupply Network
2. Non-divisional Missiles and Rockets
3. Nuclear Warheads (Type and Yield)
4. Chemical Munitions (Type and Agent)
5. Civilian Population (Size and Location)

SLIDE 11-4 CHEMICAL MODEL

OPERATIONS

- **Considers Employment Doctrine, Weapon Systems and Targets**
- **Assigns Chemical Weapons to Targets**
- **Assesses Personnel Casualties and Loss of Unit Effectiveness**

EMPLOYMENT DOCTRINE CHARACTERIZED BY

- **Employment Levels**
- **Employment Stimuli**
- **User Specified Parameters for Weapon Employment**
- **Internal Decision Rules for Weapon Employment**

TARGETS CONSIDERED

- **Subunits of Divisions, Their Personnel and Equipment**
- **Personnel Manning Nuclear and Chemical Delivery Systems**
- **Tactical Airbases, Their Personnel and Aircraft**
- **Supply Depots and Their Operating Personnel**

TARGET DAMAGE IS A FUNCTION OF

- **Size of Target and Personnel Protection**
- **Fill Weight and Delivery Mode of Chemical Munition Used**
- **Windspeed, Atmospheric Conditions, Height of Release**
- **Type and Toxicity of Chemical Agent Used**
- **Target Location and System Delivery Errors**

fired at. So we use the idea of some employment level to characterize or to control the number of munitions of a certain type that are going to be fired in a given situation.

Now, to relate this to the nuclear concept, we've made these two models parallel in structure so that when I speak of an employment level in the chemical model it's very similar to what I might call an escalation state, or an escalation level, in the nuclear model. This means that the user has to specify certain descriptions for a given situation. Let's say he wants to talk about the lowest level of escalation; at this level, the user must portray the kinds of targets he's willing to fire at, given that they're detected, where in the battlefield he will fire those weapons and what yields would be used — or in the chemical sense, what kinds of chemical agents would he fire. He uses these types of descriptors to specify three or four levels of employment. The conventional model then moves along until it hits a threshold value which then amounts to saying "I'm in a position now where I'd be willing to fire a chemical or nuclear weapon at targets, and I know the kind of tactical situation I'm in that gives rise to a certain employment level." At that point, you can then have a certain number of weapons released to fire at that level.

SLIDE 11-5 CHEMICAL WEAPON USAGE

DETERMINED BY

- **EMPLOYMENT STIMULI**
  - **- A PREPLANNED DECISION TO ATTACK**
  - **- IN RESPONSE TO SPECIFIC TACTICAL CONDITIONS**
    - **- LOSS OF SPECIFIC RESOURCES**
    - **- HIGH RATE OF MOVEMENT BY ATTACKER**
    - **- DISTANCE ADVANCED BY ATTACKER**
    - **- INABILITY TO MEET ATTACK OBJECTIVES**
  - **- IN RESPONSE TO THE ENEMY'S USE OF CHEMICAL OR NUCLEAR WEAPONS**
- **EMPLOYMENT LEVELS**
  - **- WHAT TARGET TYPES ARE ALLOWABLE**
  - **- WHAT FRACTION OF TARGETS SHOULD BE ATTACKED**
  - **- TO WHAT DEPTH WILL THE ATTACK BE CARRIED**
- **AVAILABILITY OF RESOURCES**
  - **- DELIVERY SYSTEMS**
  - **- CHEMICAL AGENTS AND DISSEMINATION MODE**

The stimuli that would give rise to firing these weapons are shown in Slide 11-5. Maybe you want to consider that it might be a preemptory strike, or a preplanned decision to attack with nuclear weapons. That's an obvious mode of attack, and the model is capable of accepting that preplanned strike at any point in the nuclear or chemical calculation.

Certain tactical situations might cause one side or the other in the conflict to use these particular munitions. If you're low on specific resources you might find that if you get below a certain level you don't have the capability to retaliate. So then you will fire before you get to that level. The other two items — the high rate of movement by the attacker, or a certain total distance advanced by the attacking force might give rise to the use of weapons. This could also happen if the attacker is not meeting certain specific attack objectives that we would like to have in the time that his attack plan called for.

There is a response time that for the delivery of these weapons encompasses delays for target acquisition, command and control, and delays associated with weapon stimuli. You can enlarge that value to incorporate a delay of almost any nature.

Now, can we incorporate a political decision kind of delay? That would have to be incorporated in the input value of response time, and we don't particularly play a separate political delay time, but it could be incorporated in one of the other inputs.

SLIDE 11-6 COMPONENTS OF CHEMICAL WEAPON ASSESSMENT

- **TARGET LOCATION AND SYSTEM DELIVERY ERRORS**
- **WINDSPEED, ATMOSPHERIC STABILITY, AGENT RELEASE HEIGHT**
- **DIFFUSION AND SETTLING OF CHEMICAL AGENT**
- **AGENT TOXICITY, PROTECTIVE ENVIRONMENTS, CHEMICAL PROTECTION**
- **LOSS OF PERSONNEL EFFICIENCY FROM CHEMICAL PROTECTION**
- **EQUIPMENT CONTAMINATION AND DECONTAMINATION**
- **RAPID ASSESSMENT METHODOLOGY**

The kinds of things that will most affect the assessments of the use of chemical weapons are shown here (Slide 11-6). Certainly, target location areas are very critical in both nuclear and chemical assessment, but more specifically critical are environmental conditions like wind speed and atmospheric stability, which means whether you are operating in an inversion, a neutral, or lapsed condition of atmospheric stability. This gives rise to different kinds of diffusion of the gases that are contained in these chemical weapons.

One of the most important things affecting personnel efficiency in a chemical environment is the necessity of having to be in a chemical protective garment for very long. There have been kill tests run showing chemical protective gear of different countries and how they will affect personal efficiency. The results showed that when someone is forced into these garments for very long he can regain combat effectiveness in a period of time. So we try to portray units being in that condition to see what effect it has on the engagement.

Question: What about the morale effects of gas? No one has seen anything of this nature for over 60 years. This might be the biggest variable. Has any work been done on that?

Dr. Kerlin: We certainly don't consider it. And I don't believe the people that I've talked to at Aberdeen and at Edgewood have done any work on it either.

Question: But the effects on the first effective use of gas in World War I were catastrophic, and it was primarily a morale effect.

Dr. Kerlin: You mean in terms of forces just getting up and leaving the foxholes?

Question: Yes

Dr. Kerlin: Well, it's sort of the same way with this personnel protection. They say that people can become acclimated to wearing these protective garments, that after experience you can wear them all day and not feel the effects of them. But I think the first time you put one of those garments on and try to run up a hill you don't reach the top. But after a number of times you may. We may be considering that most of our troops are "trained", or have the training in a chemical environment. The same question can be asked about nuclear; what are the psychological factors in a nuclear environment? I don't know how to measure those.

Question: True, but the Russians would be more likely to use gas than nuclear weapons because there's less danger of escalation to a strategic conflict.

Dr. Kerlin: Well, I don't know if the effects are more catastrophic though on people, they may be, but I don't know of anyone who's done work in that area.

Question: I'm intrigued by the fact that you've got a low resolution model and yet you don't use target acqusition for conventional weapons. Why did you feel it was essential to use them for tactical nuclear weapons?

Dr. Kerlin: Well, I guess because everybody else has. But, more seriously, I think that the ability to take out a subunit with maybe one, or a few rounds of chemical or nuclear weapons, since those resources are more scarce than, say, conventional weapons, may be an argument for being able to portray what you're actually firing at more specifically.

Question: I see that, but it seems to me that if you were really trying for a low resolution model you might try to aggregate in the target acquisition area as well.

Dr. Kerlin: That's correct. One comment I was going to make is that by going to that target acquisition model it has almost doubled our running time. When we don't play that model, our running time is half of that.

Question: Was there a serious debate about this, or not?

Dr. Kerlin: I guess there wasn't, although initially we talked a little bit about how much we should go into target acquisition. But we really did not have too many good ideas about alternative approaches. Nevertheless, it's a good point to consider for further work.

Yes sir?

Dr. Huber: Actually, I wanted to make the same point. Unless you attempt somehow to trade off your target acquisition systems or sensors, versus another resource, it doesn't seem to make much sense to have them in there. You can't assume the targets are available. I suppose, on the other hand, if you recognize that nuclear weapons are a scarce resource, you can't waste any of them. You have to have targets for them and you can't just assume the targets are available. You must go into a model that tells you what targets are available and how good your sensor systems really are.

Dr. Kerlin: That's right. I didn't bring it out, but we do play attrition of these sensors throughout both conventional and nuclear warfare. I mean, the reconaissance aircraft are being killed in a conventional role, and the ground sensors are being killed in conventional attacks. They aren't considered to be safe in the model and when destroyed they don't give us back the information that they might gather. So when attrition is going on, we just aren't getting output from the units that are affected.

Dr. Kerlin: Yes sir?

Question: Concerning your measurement of the loss of people and equipment, when you enter a nuclear phase do you in some way change the attrition rate to reflect that a nuclear deployment posture would be adopted or anything like that?

Dr. Kerlin: The rates are not changed so much, but certainly the division deployment on the ground changes. The divisions in each of the battle areas are on a frontage and depth that is set for conventional combat, a frontage and depth that is nuclear prepared, and a full nuclear deployment. So we have three different deployments for each type of division. This, of course, gives rise to more space between subunits and therefore fewer bonus effects, let's say, when nuclear weapons are fired at them. But the conventional attrition is no different when they're in this larger posture.

Dr. Bracken: Next, I'd like to introduce Colonel Carl Hess. When the organizing committee for this symposium was talking about experience with interfacing ground air and sea war combat models, the only person that we knew who had ever done this was Carl Hess. Carl has a doctorate in operations research from the University of Michigan and he's been an analyst in the Pentagon for a number of years, first with SAGA and now with I&L. Carl was the project leader for the CAPSTONE/CAPLOC series of studies and this is what he'll talk to us about today.

12 — Models Experience with Interfacing Ground/Air and Sea War Combat

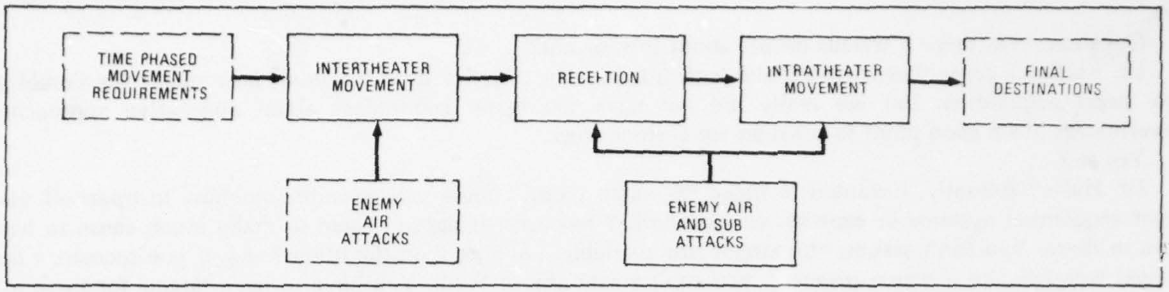
COL. CARL HESS, USA
OASD (Material Acquisition and Logistics)

The heroic employment of models in series and in parallel to solve a difficult logistic throughput problem

Col. Hess: Jerry asked me to talk about my experience in interfacing models, and those are the three key words — experience, interfacing, and models. You can strike out all other adjectives because I haven't worked on a theater level model like you have been talking about today, but I think that what I have been working on is worthy of talking about as far as interfacing is concerned.

The work I'll discuss was performed in SAGA from 1972 through 1975. I worked on the CAPFORCE, CAPLOC I and II and the CAPSTONE studies, and what I'd like to talk about today is based on those studies.

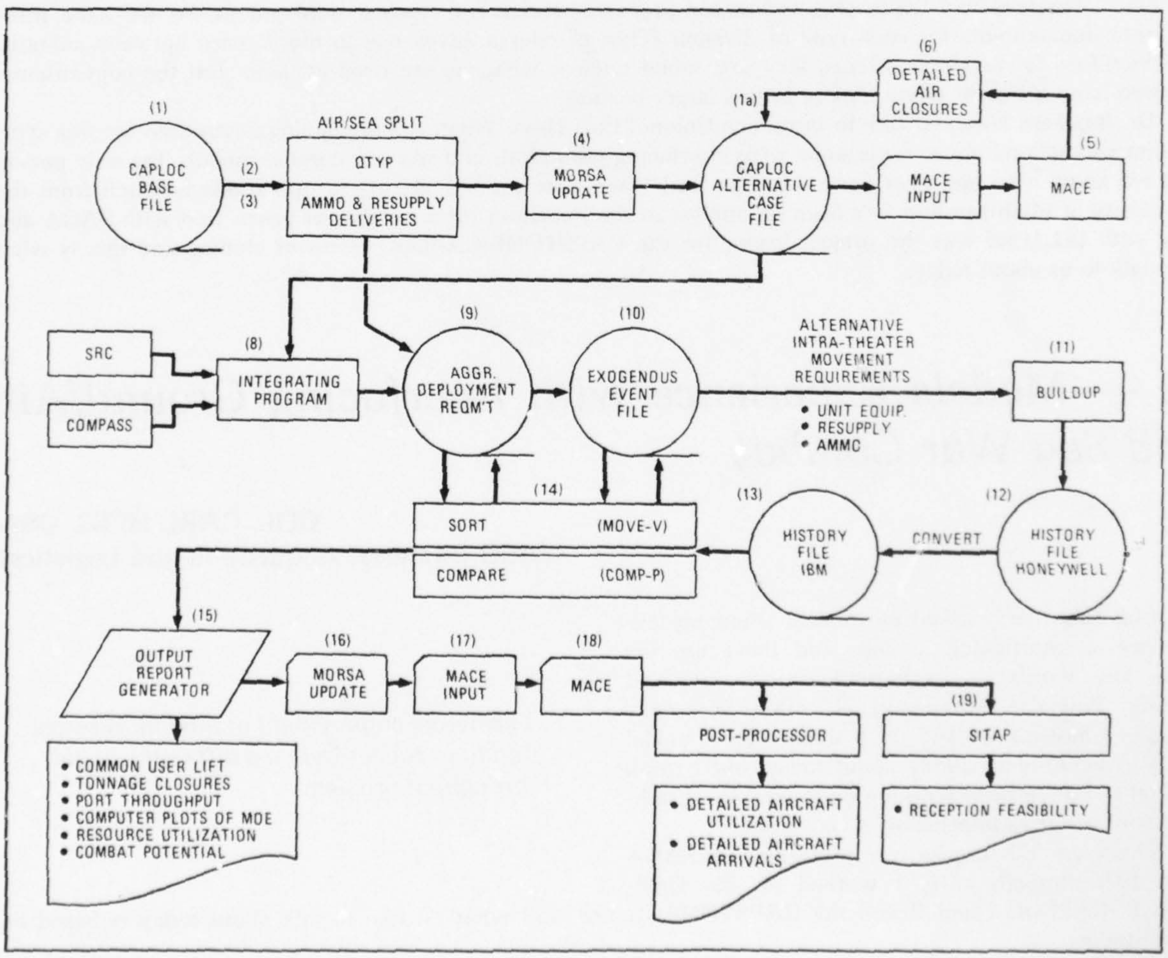
SLIDE 12-1 REINFORCEMENT SYSTEM FOR ANALYSIS



CAPFORCE, CAPLOC I & II and CAPSTONE were NATO reinforcement studies and Slide 12-1 is a schematic of the basic reinforcement system and we'll plug the various interfaces into the compartments shown in the slide. Reading from left to right on the top line, you can see how it describes the flow of reinforcement from CONUS, moving it over the Atlantic to Europe where it is received at seaports and airports, then moving it forward through the intratheater loc to its final destinations. Enroute, of course, it could be exposed to attrition by enemy action.

Now, the actual system that we used is not quite as simple as seen in (Slide 12-2), which mirrors the top line of the previous transparency. I'm not going to talk about the flow chart, but I just wanted to show you that there are a lot of interfaces when you start looking at something like this reinforcement flow.

SLIDE 12-2 FLOW CHART DEPICTING INTERRELATIONSHIP OF MODELS, ADP FILES, AND PROGRAMS



We've used multiple models for a number of reasons. The first model basically is in an interactive or iterative mode. Let me show you what I mean:

- Interactive
  - — Iterative process
- Speed
  - — Turn-around time
  - — Set up time
  - — Alternatives considered
- Parallel
  - — Reduce noise
  - — Cover gaps
  - — Check results
- Serial
  - — Output of one model is input to another model
- Problems
  - — Manual interfaces
  - — Conflicting results
  - — Timing
  - — Computer

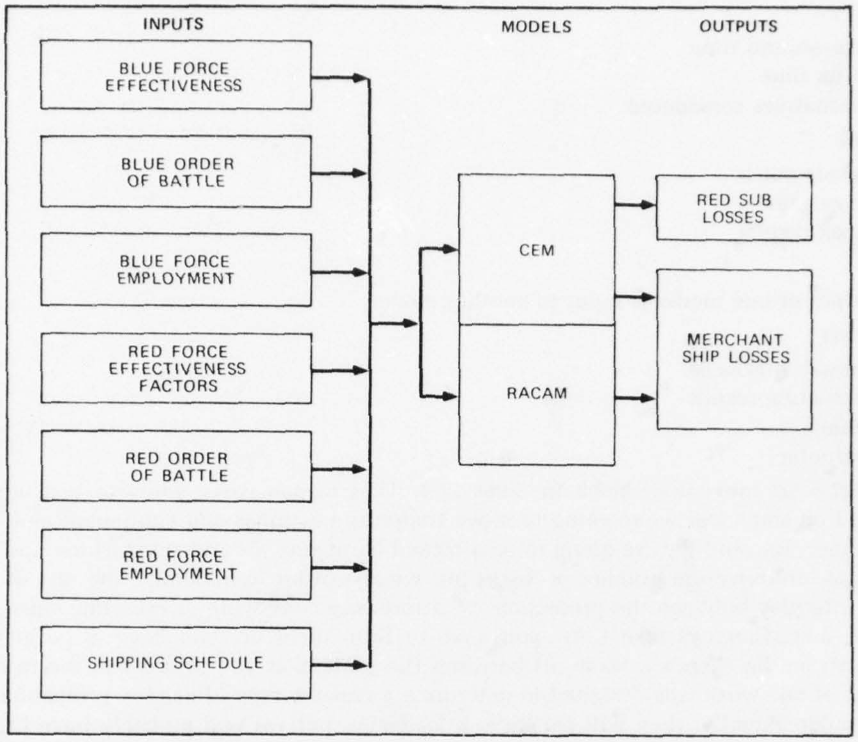
You see the intertheater movement block in Slide 12-1. That means we're going to load up airplanes and we're going to load up ships and we're going to move troops and supplies and equipment and so on from one continent to the other. Enroute they're going to be attacked by enemy air and submarines, and maybe surface vessels. We want to minimize the number of losses but we also want to maximize the rate of reinforcement, so it becomes an interplay between the protection we afford, say convoying or escorting ships, and the speed at which they advance. Convoys take time; you have to form them up, you have to go at the rate of the slowest ship, and so on. So there's a trade off between the protection you afford and the rate at which you can get there. Part of our work was designed to determine a feasible type of convoy protection scheme. If we let the ships sail independently, they will get there a lot faster, but we will probably have higher losses. Or, we could let them go slower and convoy them, diverting surface combatants from other possibly more useful missions to escort them. Exploring this interaction entailed an iterative process. We would run the strategic mobility model, and use the deployment from that model to run the sealift attrition model. That would give us certain losses. Then we'd turn around and run the strategic deployment model again with those losses, and so on. Basically, in the process, we found we'd have problems with the loss of escorts, so we couldn't run as many convoys. Then we would have to slow down our schedule, and so on. It was simply an iterative procedure back and forth between two entirely different models in order to arrive at some final judgment on what we could do.

Another use of multiple models is for the speed involved. In most of the things we did, like everybody else, we used these great big humongous models that take a lot of input time and a lot of computer time. We were always in an environment in which we didn't have enough computer time, or the models weren't working, or the computer wasn't working, or J-3 had to process a JOPS run and we couldn't get on the machine. So we would use some simple models that would mirror the same thing, in which we could get fast turnarounds. When we eventually got our consoles so that we could work up in our office space, we could run these simple models almost instantaneously, or at least we could get overnight service. And of course, we didn't have to spend a lot of time setting them up. Moreover, we could look at just a lot more alternatives. We ran these simple models to use them as a filter to get rid of the more obvious alternatives that we didn't want to look at with the more detailed complicated models that folks had a lot more faith in. But really the results were the same.

We used a lot of models in parallel. For one thing, we never really did trust the outputs. We've been talking all day about not believing the absolute numbers. Within parallel models, we'd try to use the same measures of effectiveness, or measures of goodness, if you will, we'd also try to use parallel models to cover the gaps, so that one model would do what the other one couldn't do. Thus we would check the results between them and attempt to reduce the noise. On this matter of noise reduction — I think we mentioned that this morning — there are a lot of internal processes going on in some of these models that are out of the hands of the analyst. The guy that built the model three or four years ago made certain decisions on how things were going to happen so that sometimes when you change inputs which you are certain will improve the situation it turns out that it doesn't, right? Well, sometimes using a parallel model with the same changes

will allow you to get the improvement you're looking for and thus reduce that noise. Maybe it won't, but we've used them with that effect.

SLIDE 12-3 SLOC VULNERABILITY METHODOLOGY

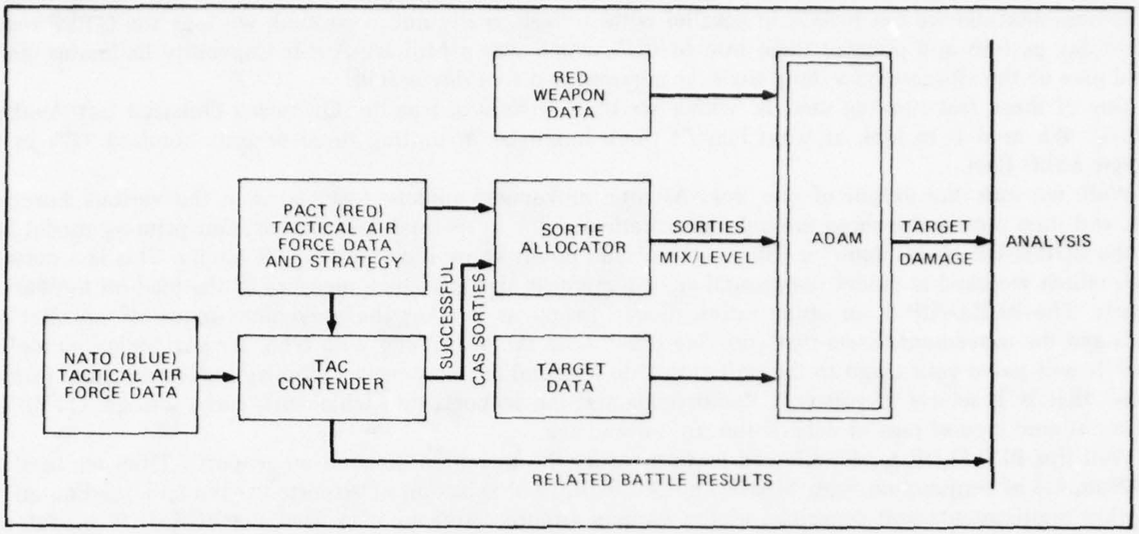


Okay, here's an example of using parallel models (Slide 12-3). This is in what you might call the "cover-the-gaps" mode. This shows the Sea Lines of Communication (SLOC) vulnerability portion of one of the studies, CAPLOC I, I think, and on the left is the list of all those good inputs that go into it. We used two models, as shown in the center. One is CEM, which is not the same CEM we've been talking about today. This CEM is a Campaign Execution Model, which I think Jerry Bracken worked on earlier. This CEM only considers convoys, and we were trying to evaluate the goodness of having independent sailings, and also the fact that we'd lost a lot of boats if they went independently. But we liked what CEM did as far as assessing the losses to the Red submarines, the losses to our escorts, and, of course, the merchant ship losses in the convoys. So, we ran it in parallel with RACAM, the Robert Ayles Campaign Model. This is a fast running model developed by a fellow in Systems Analysis. It had independent sailings in it. It didn't have nearly the fine resolution that CEM had, but we'd get the models tuned together so that we got approximately the same output from the losses from convoys, and then we'd just read the dial on independent losses. So, this is a case where we'd have to use one model to fill in the gaps created by another model.

Another use of multiple models was in a serial mode. Now, of course, you can say people always do that. They take individual weapon models and use them to build inputs for the next highest resolution model, and so on. Well, we did that too, of course, but that's not what I'm talking about here. We would use some of these big models and use the outputs from them to input into another large model.

One of our problems was to evaluate the vulnerability of intratheater line of communication facilities to enemy air attack. A key subproblem was to determine how many resources the enemy could afford to devote to these targets. Well, to determine that, we used the TAC CONTENDER model as shown in the lower left of Slide 12-4 to simulate the air battle. We took the output and used it in our Air Damage Assessment Model, or ADAM, which assessed the damage the surviving aircraft could do. We probably misused TAC CONTENDER here, I'm not saying that we didn't. You know, when you're down in the trenches and you've got a job to do, you've got a set of tools, which are the models, and you've got a set of folks who run the tools. So you try to get from the beginning to the end of the study using what you've got. So, we had this tool, TAC CONTENDER, and we jimmied it around, letting it run in a free mode for about three days — three combat days not computer days. If the forces are fairly even, TAC CONTENDER, will fight the air battle first and try and

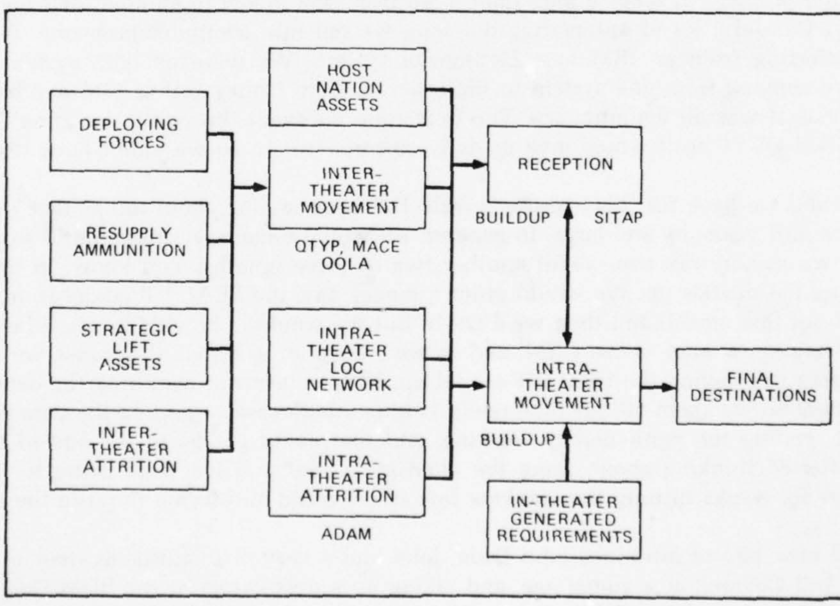
SLIDE 12-4 STRUCTURE OF THE ANALYSIS



win air superiority. So we let it try and win air superiority for three days. Then we stopped the model and diverted certain Red resources, air resources, to attacking land combat, targets, such as prepositioned stocks or port facilities and so on. Then we started TAC CONTENDER again to find out what happened to the Red battle fleet after that happened. That's what that last line is in Slide 12-4 — related battle results.

We also got out the number of successful sorties, combat air support sorties, or whatever they call them in the TAC CONTENDER literature, which we then could use to apply in ADAM to measure the damage to other facilities, which is what we were really interested in measuring.

SLIDE 12-5 FUNCTIONAL AREAS AND ASSOCIATED MAJOR ELEMENTS OF THE ANALYTICAL SYSTEM (U)



Slide 12-5 is another diagram of the earlier reinforcement system that I showed you. As a matter of fact, it pretty well mirrors the computer flow chart I showed earlier (Slide 12-1). Here we used multiple models in several cases. For the intertheater movement we used a QTYP* model, basically for the strategic deployment.

---

*It was subsequently revealed by Prof. M. Sovereign during the panel discussion that he, indeed, was the father of the QTYP model and that the acronym stood from "Quicker Than a Yellow Pad."

QTYP at that time would give us daily closures by sea, but air closures by every five-day period. Well, that wasn't good enough because part of our problem was to assess the reception capability of the various airfields on the continent. So we ran MACE in parallel with it. Well, really not in parallel; we took the QTYP results by five-day periods and plugged them into MACE, which was a Military Airlift Capability Estimater, and it would give us the closures on a daily basis, as opposed to a five-day period.

One of these fast running models, which we built ourselves, was the Oversized Outsized Lift Analyzer (OOLA). We used it to look at what benefit there might be in putting three or four hundred 747s in the strategic airlift fleet.

Well, we took the output of our cross-Atlantic movement models, received it in the various European ports, and then moved the cargo through the intratheater loc to its final destinations. Our primary model here was the BUILD-UP model done by GRC; some of you might know it as the GREER model. This is a network model which we used to model the ground environment. It also gave us a measure of the load on the various seaports. The BUILD-UP is an optimization model insofar as it takes the movement requirements that you give it and the movement assets that you give it and does the best it can with what it's got, including looking ahead. It will move your cargo to the port closest to it's final destination to minimize movement time through the loc; that is, it selects, if you will, the airfields and the seaports into which this cargo will go. QTYP just used a notional type of port of debarkation for air and sea.

Well this BUILD-UP model allowed us to measure the reception demand on seaports. Then we used the SITAP model in conjunction with MACE and the BUILD-UP selection of airports to give us a reading on the reception requirements and capacities of the various airports. And we also used the BUILD-UP model, as I said, for the intratheater movement to final destinations. Here, we would degrade the various facilities in the loc or in the ports from the results we got from ADAM to try and measure the effectiveness of the enemy air attacks that we had determined earlier.

Naturally we had many problems with all this manipulation. One of the big problems, of course, was that a lot of the processing was manual so we had to punch cards and otherwise massage the output of one model to put it into another model, and that is a big pain when you're trying a lot of alternatives. The flow for this one study took about eight days in a successful run, and we never had a completely successful run.

Another problem with using more than one model is that you get conflicting results, and, so, you've got to pick and choose among them. Hopefully, try to be objective enough not to let your biases show when you determine the results.

Timing is another problem. It takes a long time to go from one model to another, and the more you can automate the better. We did a lot of automating but then we ran into computer problems. At that time the NMCSSC was transferring from an IBM to a Honeywell system. We were on both systems with various models and we were copying from one system to the other — from Honeywell to IBM and back, and it was just a mess. Of course, it was all evolutionary. The next time we went through it, we didn't have all these problems. We modified QTYP so it would give us daily closures by air, so we didn't have that problem and so on.

Now, how long did we have for this exercise? Well, I've been talking about things that I've done over a period of three and a half years, by and large. In general, we would have, say, nine months to do a study and then for credibility we usually ran over about another two or three months. You know, in these things, the big problem is getting the models up. We would order a model, say, the SEALIFT model from Jerry Bracken, and he would grind out this model, and then we'd get it. But we wouldn't be able to use it because our folks didn't know how. They knew how to use CEM, and so we'd keep on using CEM because we couldn't afford to devote the resources to bringing the SEALIFT model up, both in terms of acquiring the data base and getting the people trained. So we spent all our time really getting a base case — getting the data base up, getting the models running, getting the methodology working, and that about got us to the end of the study time allotted. Then we started thinking about doing the alternatives we said we were going to look at. So we would spend the last six weeks turning the crank as fast as we could and trying to write the study report at the same time.

Dr. Bracken: I'd now like to introduce John Bode. John had a very distinguished career in the Air Force before retiring as a full Colonel, at a young age, and taking up a new career at the BDM Corporation where he's been leading groups of model builders and analysts. We were looking for someone to address a tough problem of aggregation and resolution in theater level models, and we thought of the BDM and Vector competition for the CASM contract. So it was natural to ask John to give us a talk.

13 — Problems of Aggregation and Resolution in Theater-Level Models

MR. JOHN BODE
BDM Corporation

Coming to grips with problem-dependent issues of aggregation and resolution in theater models

Mr. Bode: Our topic is aggregation, which Jerry (Bracken) points out is a very difficult problem. I decided I had a choice of either saying a lot less than I wanted to say, or trying to compact it into a dual slide briefing. So, I'm going to try on two slides. If you don't get it don't worry about it. It turns out that I don't think that we have done the theoretical work that we should have been doing in modeling work in the past few years. I agree with General Welch's comments this morning. Partly as a result of some of the jobs we've won and lost both recently, we have stepped back and tried to take a look at the underlying theoretical structure in military combat modeling and see how it relates to other areas of simulation in modeling that have been done in the past.

SLIDE 13-1 WHAT IS AGGREGATION?

WHAT IS AGGREGATION?

AGGREGATION IS THE "LUMPING TOGETHER" OF SEVERAL INDIVIDUAL THINGS INTO A COMPOSITE THING WHICH IS THEN USED TO COLLECTIVELY REPRESENT THE INDIVIDUALS

AGGREGATION CAN BE VIEWED AS:

- A TRANSITION FROM INDIVIDUAL (OR MICRO) PROPERTIES TO ENSEMBLE (OR MACRO) PROPERTIES (NATURAL SCIENCES)
- A SELECTIVE ENCODING OF KEY INFORMATION WHICH "SUMMARIZES" A GROUP OF INDIVIDUALS (COMMUNICATIONS SCIENCES)
- A MANY-TO-ONE REPRESENTATION OF INDIVIDUALS IN ONE SYSTEM BY INDIVIDUALS IN A COARSER AND LESS COMPLEX SYSTEM (SYSTEMS SCIENCES)

As far as aggregation is concerned, it can be looked at in many different ways (Slide 13-1). Some of them are different from the way that we normally do it in combat models, and we just picked up a couple of examples here. In thermodynamics you can contrast microthermodynamics and statistical mechanics to macrothermodynamics. I think there are some lessons there — I'll leave that to you to think about. Coding theory, General Welch mentioned that point this morning. The F-15 is a code for a massive set of data meaning one thing to one person and another to another. But you can set up codes like that. We do it all the time in communications theory. In economics and social theory you aggregate by groups, actuarial tables aggregate by classes of people, and all those sorts of things. So, you get all kinds of coding structures that we generally can call aggregation.

SLIDE 13-2 PROPERTIES OF AGGREGATION

As far as we're concerned here, the key point of aggregation (Slide 13-2) is that it is information reducing. It does not carry the information. It carries the code, and maybe it can be tracked back and maybe not. But in a sense it is irreversible. Once you make the aggregation, then you can deal with the aggregate properties, but you cannot create the elements in detail any longer. So, in military models, when you ask for aggregations don't expect to be able to get detailed disaggregation that carries along all the properties of the detail because it's not theoretically possible, and I think that's important.

SLIDE 13-3 KEY PROBLEM FACING FORCE DEVELOPERS & PLANNERS

Now, why do we care about aggregation? Well, I think the nature of our problem has changed (Slide 13-3). The key word in here is leverage. We've got to look at new doctrines and tactics, new weapons of mass destruction, problems of the transparent battlefield. We're not dealing with force problems any more in the main. Back in the days of ATLAS, it was a question of how many divisions should we have in the field. When we were working TAC CONTENDER, it was how big should the air force be, questions of things like shelters. Now, the questions are becoming more sophisticated (Slide 13-4). What kind of command and control structure do we have to have. Where are the key vulnerabilities in Soviet operations? How sensitive are we in certain kinds of logistics? How soon do we have to get the nuclear weapons out of the storage pile? How do the nuclear weapons and the tactical nuclear force impact on the theater? It turns out all of these things are very important, but they're all leverage questions, and not macro problems any more. So, in a

SLIDE 13-4 ASSOCIATED AGGREGATION PROBLEM

ASSOCIATED AGGREGATION PROBLEM

LEVERAGE RESIDES IN THE FINE STRUCTURE

- MANEUVER
  - BREAKTHROUGHS
  - ENVELOPMENTS
  - MEETING ENGAGEMENTS
  - DARING THRUSTS
  - INTERMINGLING
- ATTRITION
  - STRUCTURAL DISTRIBUTION
  - FUNCTIONAL DISTRIBUTION
- COMMAND, CONTROL, AND INTELLIGENCE (C^2I) PROCESSES

THESE ARE NOT NEW PROBLEMS, BUT THE EVOLVING TRENDS IN THEATER LEVEL WARFARE ARE MAKING THEM MORE AND MORE CRITICAL.

sense, the macro aggregation that we've used in the past probably isn't adequate. One area in which that shows up more than any place else is in the area of maneuver. In the old days, we talked about massive attacks, and linear defenses. Now, we've rammed those battlefields with Phil Karber's pot pourri — it could be a "daring thrust", and it's a completely different ballgame. You simply cannot represent this process the same way you can the classic breakthrough (Slide 13-5).

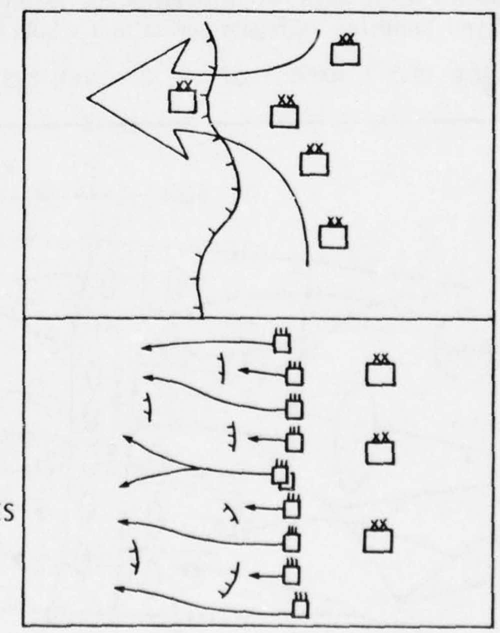
SLIDE 13-5 NEW EMPHASIS ON PRE-EMPTIVE MANEUVER

CLASSIC BREAKTHROUGH

- LINEAR DEFENSE
- MASS FOR ATTACK
- CONCENTRATED FIREPOWER
- DIVISION/ARMY LEVEL

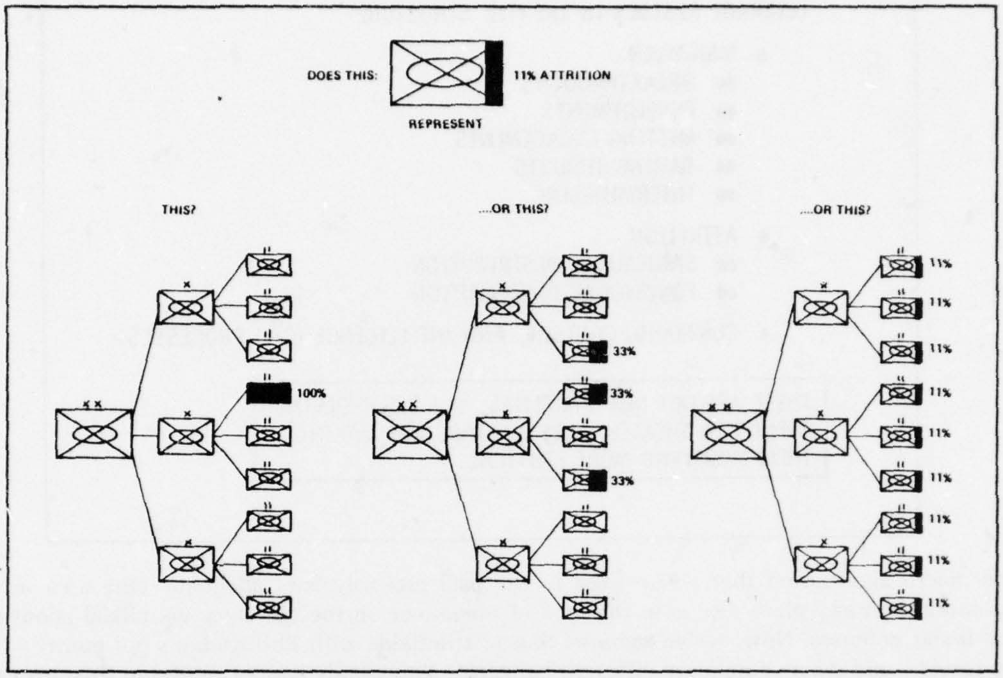
"DARING THRUSTS"

- GRANULAR DEFENSE
- MEETING ENGAGEMENT
- MANEUVER ON MULTIPLE AXES
- REGIMENTAL LEVEL



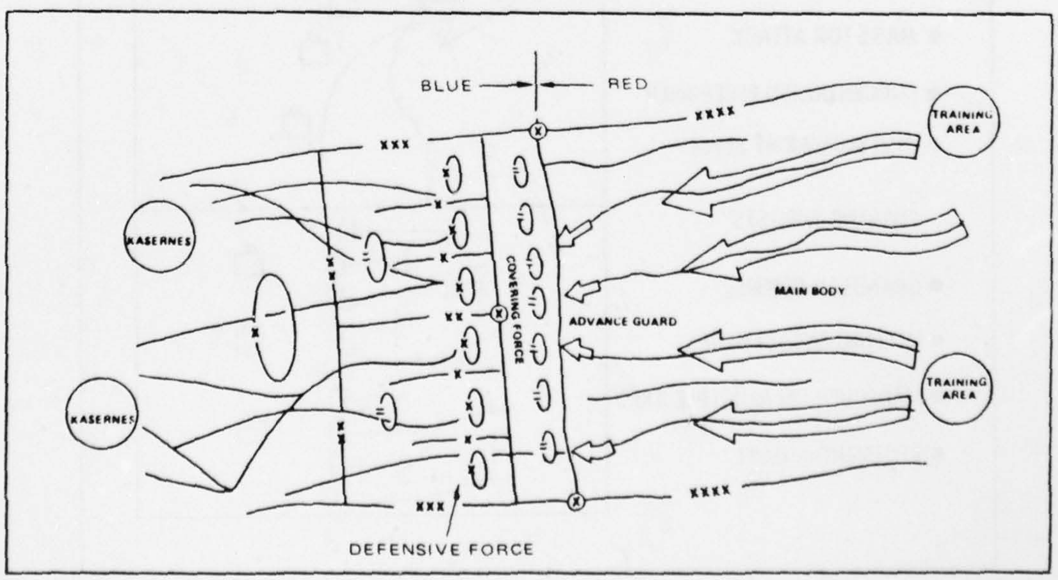
The same thing is happening in attrition and command and control. None of these problems is new, but the trends in military combat are making these problems more crucial, so we have to find some new ways to deal with them.

SLIDE 13-6 IMPACT OF STRUCTURAL DISTRIBUTION OF ATTRITION



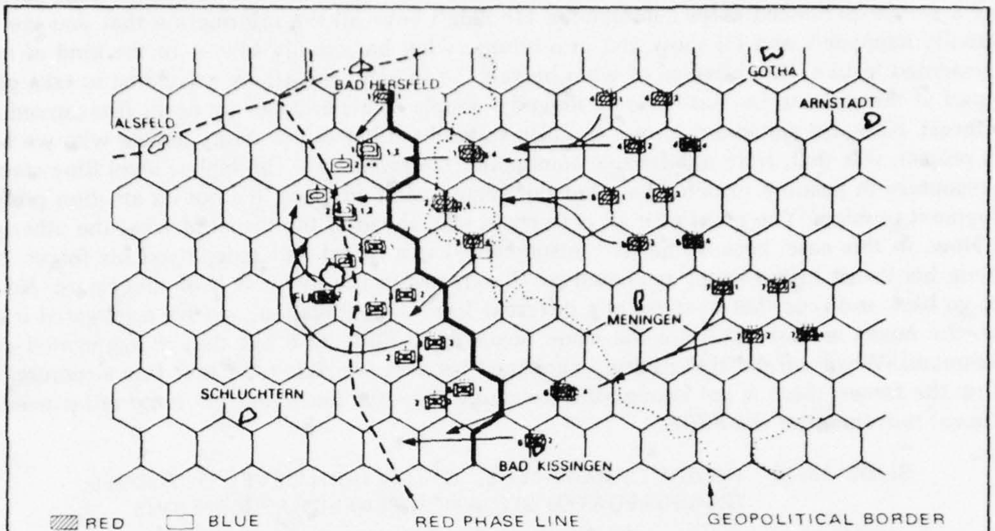
Let's take a look at just simple attrition (Slide 13-6). Take an aggregate representation of attrition of, say 10-11%, and then look at the different ways you can represent that 10% at different levels of aggregation. Obviously, those things are not the same, and the difference makes the difference. Now, here's another thing that I haven't heard mentioned lately — that is the functional distribution of these processes (Slide 13-7). Here we have a case where we have two different kinds of forces involved, the covering force and the defensive force. Now, if you lump the aggregate across that whole structure, it makes an awful lot of difference

SLIDE 13-7 IMPACT OF FUNCTIONAL DISTRIBUTION OF ATTRITION



whether or not 10% attrition is taken out in that covering force or whether it's taken out in the defensive force. It could make the difference between a successful penetration and an unsuccessful penetration.

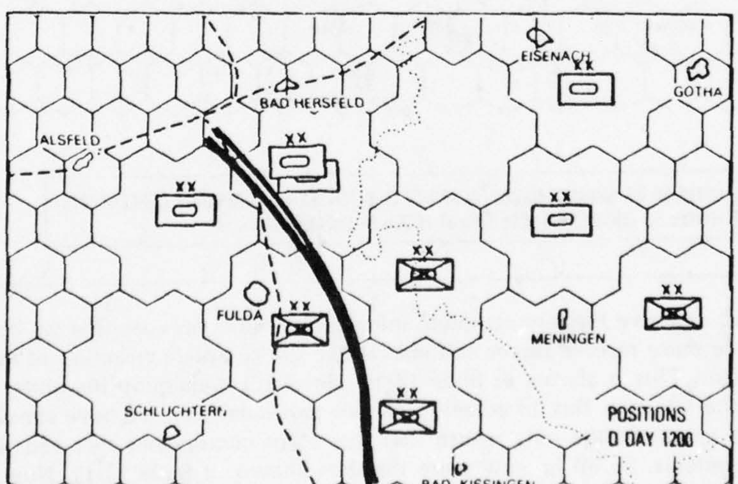
SLIDE 13-8 IMPACT OF LOW-LEVEL C^2I ON HIGH-LEVEL OUTCOMES



IN THIS PARTICULAR OUTCOME OF A MAN-IN-THE-LOOP EXPERIMENT, OVERREACTION TO THE PERCEIVED THREAT IN THE BAD HERSFELD REGION OPENED UP THE MENINGEN-BAD KISSINGEN-SCHLUCHTERN REGION TO AN ENEMY PENETRATION

I want to show you a couple of sequences that we've gotten from some recent models that we've been running (Slide 13-8). We are doing a lot of modeling development, I might add. Most of it accidental, I guess,

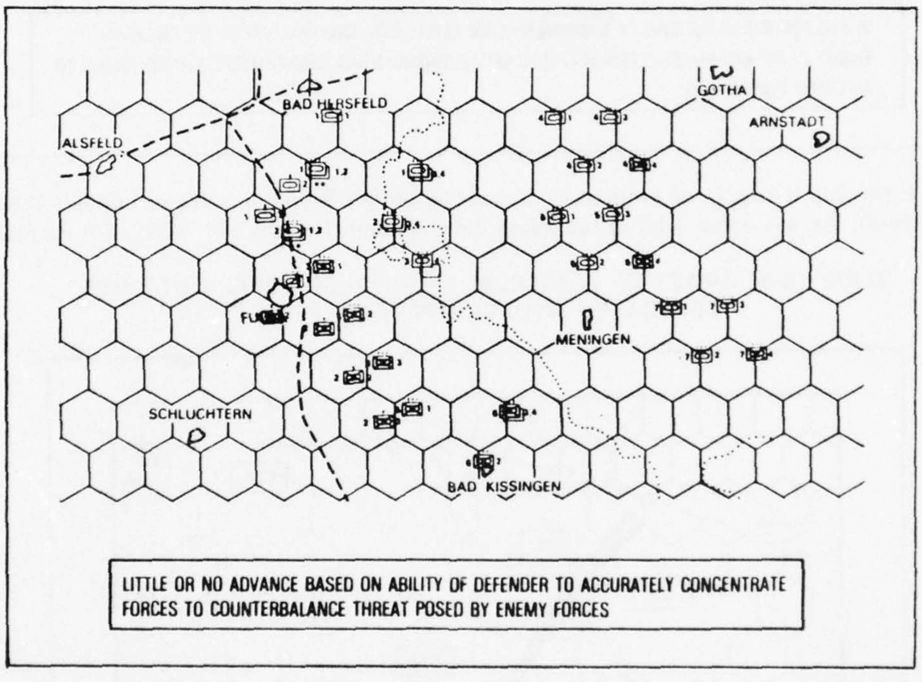
SLIDE 13-9 IMPACT OF LOW-LEVEL C^2I ON HIGH-LEVEL OUTCOMES: AGGREGATED SITUATION/PERFECT PERCEPTION



LITTLE OR NO ADVANCE BASED ON RELATIVE STRENGTHS AND ATTACKER-DEFENDER RATIOS OF FORCES IN CONTACT

but this particular case is a run on a hexagon system of one of the models that we are developing at present. Very briefly, let me point out some things. There is the geopolitical border, there is the Red phase line for an attack, there are two Blue divisions each arrayed with three brigades, and there is a Soviet attack, in which the Red units come up against the Blue forces as shown. So these hexes are essentially brigade size in this particular case. Now, this is an example of a man in the loop exercise that we have run. This particular case was run by a very experienced corps commander. He didn't have all the information that you see here. This is what actually happened, and I'll show you in a minute what he actually saw — or the kind of information that was presented to him. But, because of what he saw, he thought the attack was going to take place across the upper part of the slide, to the north. So he moved a couple of his brigades up north (the curved arrows) to block this thrust. Now, his statement to me, and also stated by other senior Army people who we have working on this project, was that, from a defensive standpoint, the problem of the higher level Blue commander is to get his resources in position to defend so that the enemy can't advance. It's not an attrition problem, it's a force deployment problem. The point of it all is to try to stay ahead of the thrust of what the other guy is trying to do. Now, in this case, because he had misperceived this thrust and redeployed his forces, the enemy actually made his thrust to the south as shown by the arrows in the lower half of this figure. Now, I asked my guys to go back and redo that exercise at a different level of aggregation, and we aggregated it at division level. Now, the hexes in Slide 13-9 are the same hexes as in Slide 13-8 but they're aggregated at a higher level of command. When we did that nothing happened because we didn't get that fine structure. So, if you just count up the forces, there is not enough difference in this situation to get the force ratios needed to produce significant movement of the FEBA.

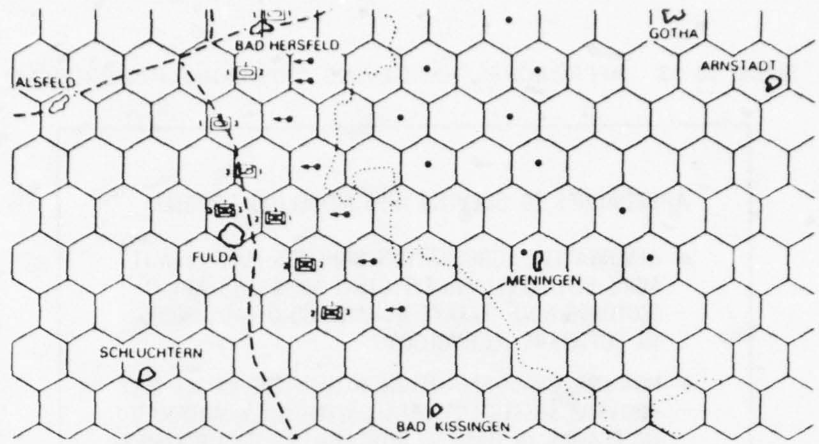
SLIDE 13-10 IMPACT OF LOW-LEVEL C^2I ON HIGH-LEVEL OUTCOMES: DISAGGREGATED SITUATION/PERFECT PERCEPTIONS



Now, if you go back and give the guy complete information, (another case that we ran), where the corps commander is able to see these reserve forces and able to see the complete structure of the force that he was faced with, he did it right. This is shown in Slide 13-10. He didn't maldeploy his units, and the Red forces still didn't get through the advance. But, in actuality, in this particular run, we have simulated Comint, Sigint Sensors flying over this area, and the only return that this corps commander received was the returns that came from the sensor systems. So all he saw were the dots shown in Slide 13-11. Now, he got these from direct acquisition resources engaging us to the forward brigades. He got these from guard rail types of sensors. So, instead of seeing the total structure of Slide 13-10 he only saw the dot structure (Slide 13-11).

This is just a snapshot of something that developed over time, but obviously it's force-heavy to the north. He thought that was the only force he was dealing with and moved his units up and he lost. I only show this as an example of the point I'm trying to make — that the detail is the place where you're going to catch the

SLIDE 13-11 IMPACT OF LOW-LEVEL C^2I ON HIGH-LEVEL OUTCOMES: DISAGGREGATED SITUATION/IMPERFECT PERCEPTIONS



IMPERFECT PERCEPTIONS PREVENT DEFENDER FROM "SEEING" FULL EXTENT OF ENEMY FORCES IN MENINGEN BAD KISSINGEN REGION

leverage for some of the kinds of problems that we're really working about today. That's the problem for us from an aggregation standpoint.

To work these problems even at the theater level, what this means, in our opinion is that we're going to have to have organizational resolution (Slide 13-12). We're quite sure we need to get down to battalion, and it's certainly true in nuclear operations, and probably true in conventional. We'd like to get down to company resolution. We certainly have to have sufficient regional discrimination to discriminate one type of target from another to spread out those command and control structures.

SLIDE 13-12 CONSEQUENCES FOR THEATER-LEVEL AGGREGATION REQUIREMENTS

CONSEQUENCES FOR THEATER LEVEL AGGREGATION REQUIREMENTS

A HIGH LEVEL OF RESOLUTION IS NEEDED IN TWO TRADITIONAL DIMENSIONS OF AGGREGATION:

- ORGANIZATIONAL - AT LEAST TO BATTALION, PREFERABLY TO COMPANY
- REGIONAL - TWO DIMENSIONAL AND SUFFICIENTLY FINE GRAINED TO PORTRAY ORGANIZATIONAL RESOLUTION

MODELING PROBLEM: HOW TO ACHIEVE THIS LEVEL OF DISAGGREGATION WITHIN REASONABLE DATA REQUIREMENTS, SET-UP TIME REQUIREMENTS, AND RUN-TIME REQUIREMENTS?

Now, how are we going to do that and fit it all into a finite machine in finite time? Well, our feeling is you have to do something different. One way is to look at new aggregation schemes (Slide 13-13), but aggregation really drives the structure of the models, the structure of the problem you're trying to work, and there's also a possibility that you can use nonuniform resolutions (Slide 13-14). We've tried both of these avenues.

SLIDE 13-13 APPROACHES TO SOLVING THIS MODELING PROBLEM

APPROACHES TO SOLVING THIS MODELING PROBLEM

- ALTERNATIVE AGGREGATION SCHEMES MAY PERMIT MORE EFFECTIVE AND EFFICIENT MODELING ARCHITECTURES ABLE TO TAKE ADVANTAGE OF ADVANCES IN SOFTWARE TECHNOLOGY.
- NON-UNIFORM RESOLUTION WITHIN THE MODEL MAY PROVIDE SUFFICIENT DETAIL WITH ONLY MODERATE INCREASES IN DATA REQUIREMENTS, SET-UP TIMES, AND RUN TIMES.

SLIDE 13-14 AGGREGATION AND ARCHITECTURE

AGGREGATION AND ARCHITECTURE

A MODEL'S AGGREGATION SCHEME DETERMINES

- WHAT TYPES OF ENTITIES MUST BE REPRESENTED
- HOW--AND TO WHAT DEGREE--THESE ENTITIES MUST BE DISTINGUISHED

AS SUCH, AN AGGREGATION SCHEME IS A FUNDAMENTAL DRIVER OF MODEL ARCHITECTURE

Here are some of the ways you can aggregate (Slide 13-15). There are really two components here, one we call characteristics. We talked about functions, goals, positions, resources, capabilities. A lot of our classic work has been done on capabilities that we talked about, firepower potential and that sort of thing. Position people like to aggregate by location. Somebody talked about aggregating target acquisition a little bit ago. I can't understand how you can do that. I have no concept of what the aggregate of target acquisition is. I can think of analytic expressions that represent target acquisition, but I can't think of a way to aggregate the process. But then you've also got relationships that you have to worry about, and one of the places where we got theoretically prostrated in VECTOR, which we thought was very interesting, in the way that combat was aggregated. The VECTOR model essentially drew circles around the interacting players and defined simultaneous differential equations that described the interaction. We didn't do that — we did something else, which I'll show you in a minute. It's not important that there's a difference of opinion there. The important thing is that there are a lot of different ways to treat interactions. What you do is going to drive the validity in the drafting of the model at some point.

SLIDE 13-15 ALTERNATIVE AGGREGATION SCHEMES, I
CHARACTERISTICS OF FORCE ELEMENTS

• FUNCTION	GENERAL ROLE AND MISSION OF THE ELEMENT
• GOALS	SPECIFIC OPERATIONAL EMPLOYMENT OF THE ELEMENT'S FUNCTION
• POSITION	REGION OCCUPIED BY THE ELEMENT
• RESOURCES	MEN AND MATERIEL CONTROLLED BY THE ELEMENT
• CAPABILITIES	POTENTIALS OF THE ELEMENT FOR PERCEPTION, DECISION/CONTROL, ACTION (FIRE, MANEUVER, MOBILITY) AND REGENERATION
• VULNERABILITIES	POTENTIALS OF THE ELEMENT FOR EXPLOITATION BY ENEMY PERCEPTION AND ACTION
• PLANS	INTENDED OPERATIONAL EMPLOYMENT OF THE ELEMENT'S CAPABILITIES TO ACHIEVE ITS GOALS CONSIDERING ITS VULNERABILITIES

ALTERNATIVE AGGREGATION SCHEMES, II
RELATIONSHIPS AMONG FORCE ELEMENTS

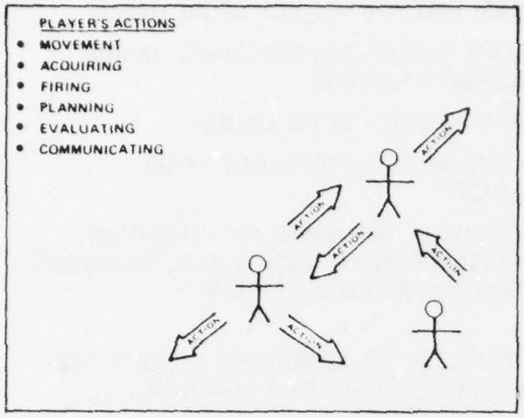
• SPATIAL	DEPLOYMENT OF ELEMENTS
• COMMAND	FORMAL CHAIN-OF-COMMAND ORGANIZATION AMONG ELEMENTS
• CONTROL/COORDINATION	OPERATIONAL ORGANIZATION OF ELEMENTS CONSIDERING ATTACHMENTS, ADJACENT ELEMENTS, ETC.
• COMMUNICATION	INFORMATION TRANSMISSION AND RECEPTION AMONG ELEMENTS
• ENGAGEMENT	FIRE AND MANEUVER AMONG OPPOSING ELEMENTS

So, what we had to do is pick up, in our way, a scheme that straddled these issues to some extent. We've aggregated around players (Slide 13-16), and Al Dobieski reminded me at lunch today that that's not a new idea for a detailed force model. They did it in CATTS, it was done in DYNTACS, CARMONETTE, the lower level resolution models. As far as I know, it hasn't been done in theater before.

However, this aggregation around players is different in the sense that we don't write simultaneous equations about sets of players, but we write the model around each individual in the process and let his perceptions and the things that happen to nim determine his reaction. We separate the mental processes from the physical things that happen.

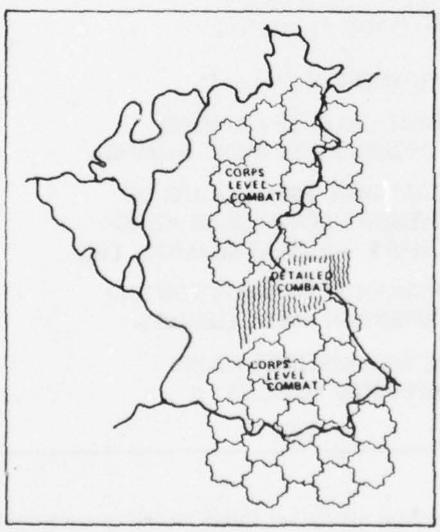
Another thing that we tried to do is look at nonuniform resolution in the theater (Slide 13-17), and I think this scheme will work. It does have some problems, but the idea is that if you have to have the detail, you certainly cannot implement detail across the theater because it won't fit in the machine — and 10^{30} is a long time — but at the same time we need that detail to be able to believe that we have represented the processes properly and that we should believe in the more aggregate things. So the problem is very highly nonlinear and it depends on the boundary conditions. Well, one way that may work to settle that boundary condition problem is to embed this detailed process in a more aggregate theater level process and only demand

SLIDE 13-16 AGGREGATION BY FORCE ELEMENT- - "PLAYER-CENTERED MODELING"



- PROVIDES A DIRECT AGGREGATION SCHEME WITH A NATURAL RESOLUTION SCHEME (COMMAND LEVEL)
- SUPPORTS MULTIPLE LEVELS OF RESOLUTION BY COMMAND LEVELS
- PERMITS ACTION-REACTION MODELING IN AN EFFICIENT SOFTWARE FRAMEWORK

SLIDE 13-17 NON-UNIFORM RESOLUTION



- PROVIDES A FINE-STRUCTURE MODEL IN A THEATER-LEVEL CONTEXT (BOUNDARY CONDITIONS)
- ESSENTIALLY IMPLEMENTS THE "HIERARCHY OF MODELS" CONCEPT WITHIN A SINGLE MODELING FRAMEWORK (ALLOWS ITERATIVE TUNING OF WHOLE MODEL)
- MODERATE INCREASES IN DATA REQUIREMENTS SET-UP TIMES, AND RUN TIMES BY CONTRAST WITH THEATERWIDE HIGH RESOLUTION
- POTENTIAL FOR FURTHER SAVINGS BY EXTENSION TO VARIABLE RESOLUTION DYNAMICALLY SELECTED BY THE MODEL (DURING EXECUTION) BASED ON WHAT IS HAPPENING.

on your periphery that the reactions be good enough to get the schedule of boundary conditions right. When I'm talking about scheduling, I mean the number of air strikes that you can expect into this area for ground support, the number of nuclear weapons that you could expect, letting the theater level commander look at the problems all the way down the front so that your reserves that are coming into this force are adequate for the total theater stress, and then zooming down in on this level to try to get the details and see if you can believe that.

Now, in principle, you can apply this detail process to any one of the plays or a number of plays that can validate what you're doing. You just can't do it everywhere at once.

So, that's sort of where we are. To summarize:

- Aggregation is information reducing and irreversible
- Aggregation schemes drive model architectures

- Both characteristics and relationships must be considered in the aggregation schemes
- Because "force multipliers" operate in the details, classical aggregation schemes are inadequate
- New theories of aggregation in modeling are being developed and tested.

I think it's well for us to remember that aggregations are going to be information-reducing and irreversible. We're not going to be able to do anything about it once we make that choice, not in a particular run, driving the model architecture. My suggestion is that anybody who wants to write a model had better work his way through this whole problem before he starts. Otherwise, as someone pointed out this morning, you can get in the middle of it somewhere and you won't be able to go any farther.

As a matter of fact, I think that is happening to us. I don't think the classical aggregation schemes that we have used will go much farther theoretically than they have right now. I think we're sort of mining dead ground, but I do think that it is possible to make some progress, and I hope — it's too bad Andy Marshall is not here — we will take up his challenge, I hope in the next couple of years we'll be showing something that does represent these processes. I believe we can do it.

Dr. Bracken: Any questions?

Dr. Kerlin: How much do you find data to be a problem in these different aggregation schemes when you go from corps level analysis in one part of the theater to very detailed levels?

Mr. Bode: Our experience has been that it's much harder to get data for the aggregate models than it is for the detailed models, because all the aggregate models require a good deal of processing, but there is a lot of data out on the detailed interaction. It's just like Seth (Bonder) pointed out this morning — that stuff has been done for years. The data are around, but if you want to abstract them and get an average kill probability, for example, like we had to do in Combat 2, again, you have to do an awful lot of work, and you have no idea whether your input's any good or not.

Dr. Kerlin: My real question is do you carry sets of data for all levels of resolution you expect to play or do you develop your higher level data from your lowest level of detail?

Mr. Bode: Well, let me answer that correctly in principle. We have developed a data base for each in the runs that we're making right now. I mean, in practice, that's what we're doing. In principle, I'd like to get to the point where we could develop analytical representations that can serve as data for the higher levels. I'd also like to add that I think Seth's (Bonder) idea of being able to take these things to analytical representations is the right theoretical approach. I think the reason we need the detail right now, though, is because we don't have any confidence that we've modeled processes properly unless we modeled them at a level of detail that's recognizable by the military commanders. It just means a very great deal to have this experienced general officer look at the thing and say, "well, yes, I think that's probably the way it really happened." If you abstract it beyond that point, it's just hard to understand what's happening yourself, and you're really asking an awful lot of the decision maker to try to make any sense out of it.

Question: What's the experience of these general officers?

Mr. Bode: The one who did this one was a very experienced corps level commander. And this particular run was viewed by several ex-theater level commanders, who had been in the Army or Air Force as recently as a couple of years ago and who had combat experience in World War II and in Vietnam.

Question: Did you consider the Vietnam combat experience applicable?

Mr. Bode: Yes, in a lot of ways we did. In terms of the resource allocation problem, certainly, but also in terms of some of the lessons that we had to relearn. The tactics were completely different, but from the corps level commander's standpoint they weren't too different, I don't believe.

Question: Do you think that interactive gaming, player gaming, plays an important role for information gathering in the decision process? We talked about the example of looking at what happens if a commander does not have good information on the enemy and he has to make estimates on the enemy's situation. There you have to look at individual or sets of individual decision correctors.

Mr. Bode: Yes, real people.

Question: Do you think that you need some amount of interactive gaming to get the data base?

Mr. Bode: Yes, in two respects. One, on the value of intelligence, is assembling that kind of support measure that is basically primary target acquisition input, and the decisions relevant to that. I don't believe we understand those processes well enough to model them without a game. Second, one of the inherent problems with this kind of model is that they're open loop models, just like real organizations are open loop, and they require the command and control structure to close the loop and keep them stable. It's impossible, in my opinion, to write a higher level control model to keep all those pieces in a stable state. So you have to have a man to look at it every now and then, replan it, and give everybody new instructions. And, in my opinion, we'll have to do that for quite awhile before we can model our way around the problems.

Dr. Bracken: The next talk will be by Dr. George Pugh. George has been writing on measures of effectiveness for military planning, and he said he's got a new set of theories which seem very interesting. George was with the Institute for Defense Analyses for many years and was one of the founders of the Lamda Corporation. Just recently he started a new company, Decision Science Applications, Inc.

14 — Outcome, Effectiveness and Decision Criteria for Combat Gaming

DR. GEORGE PUGH
Decision Science Applications, Inc.

A concept that emphasizes the role of value judgments in the modeling of decisions processes

Dr. Pugh: What I'm going to do today will be a bit of a change of pace. So far, almost all of the talks have dealt with the details and the nitty gritty of the very difficult issue of theater level war gaming. What I'm going to address are some philosophical and conceptual issues that have to do with the measurement of outcome, the definition of measures of effectiveness and the decision process itself. I'm going to talk about these things because the representation of the decision process within the combat simulations has been one of the most difficult and troubling issues that we've had to address.

To do this, I'm going to have to go back to fundamentals. I hope you'll excuse me if, in some respects, the talk seems very general and lacking in detail, but I want to cast the discussion in a simple context so that I can bring out what I think are some interesting concepts. I don't think that at the end anyone is going to have a much better idea about what specific measures of effectiveness one ought to use or what specific outcome measures one ought to use. But I do hope that you may get a little better understanding of some of the sources of difficulty in trying to select satisfactory measures. So, I'm going to embed the discussion of the various kinds of measures within a decision theory perspective.

Some of the specific measures that we're talking about are *outcome* measures that are really concerned with our judgments concerning which of a number of possible outcomes for a simulation is preferred. A second, closely related type of measure is one we call decision criteria. If you're doing decisions within the simulation, these criteria are concerned with deciding which course of action is preferred among a number of alternatives. A third type of measure is something we call measures of effectiveness. These measures are used when what you're really concerned with is some way of measuring which system or which combination of systems is best.

Obviously, in all of these cases, the fundamental purpose of the measures is to assist in some kind of overall decision process. In each case, the measures serve essentially as value criteria in making decisions. Actually, most people are quite good in using intuitive value judgments to make decisions. The problems arise when we try to deal with such judgmental values within the context of a formal analysis. Then we have trouble because our scientific training tells us that we should address the problem objectively — without resorting to subjective judgmental criteria. Consequently, there's a conflict between our intuitive way of approaching the problem and the way we think we *ought* to try to approach to problem from a scientific perspective.

It is now becoming clear to me from the work that I've been doing on decision theory that the way we use values in our common sense reasoning is actually justified by some fairly fundamental principles of cybernetic efficiency. Indeed, we could probably do a much better job in our analysis if we learned how to use some of those same general principles about judgmental values within the analysis process itself.

What I hope to be able to show today is that the judgmental values really ought to play an important role in our analysis and modeling. Moreover, by deliberately incorporating some of the subjective value criteria

within our models there's a likelihood that we would be able to build models that are considerably more flexible and more efficient than the ones that we've been dealing with in the past. There is, of course, one obvious reservation. If one is going to use models that include subjective value criteria, then it's very important that the value criteria be explicit so that people can understand the outcome within the context of the value criteria that have been used. So, what I'm going to do now is go back to fundamentals and discuss the role of values in the decision process from a decision theory perspective. Then I'm going to move on to the implications for war gaming methodology.

Some of the traditional benefits of values within value criteria is that they make it possible to decentralize the decision process and still get reasonably sensible results. It's also helpful in enabling us to make decisions in terms of intermediate outcomes when we're unable to project the outcome of the decision all the way to an ultimate outcome.

Values also serve in a very practical way as a tool of command. One of the things that a commander does in giving commands to his subordinates is to define value priorities or priorities for his course of action. The commander's action changes the value priorities by which the subordinate commanders make their decisions.

SLIDE 14-1 DEFINITIONS

VALUE:	A scalar quantity, associated with outcomes for the purpose of making decisions
VALUE FUNCTION:	A scalar function defined over the space of outcomes for the purpose of making decisions
	[typical form, $V = \sum_i f_i(u_1, u_2, \text{--} u_n)$]

If we're going to talk about values in more or less a quantitative way from a modeling perspective, we have to be considerably more formal about them than we are in our ordinary conversation. So, I have here (Slide 14-1), a microspective of what a value is. I define a value to be a scalar quantity that is associated with outcomes for the purpose of making a decision. To make decisions among a number of alternatives, we usually need a value function. A value function is a similar kind of scalar function but defined over the space of possible outcomes, so that you can assign values to the various outcomes and make choices between them.

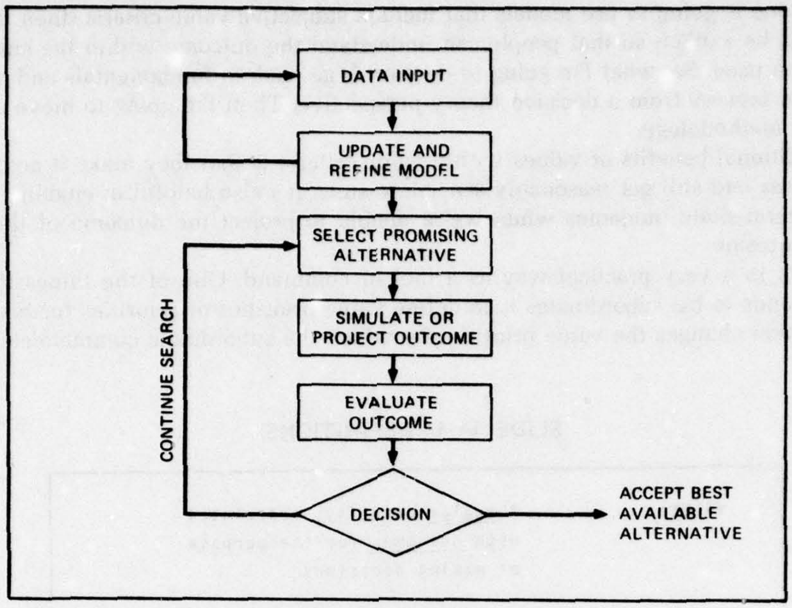
Now, the typical form of a value is essentially a summation over a number of considerations of a series of values that are somehow functions of the outcome. So it generally takes the form of a weighted sum of the number of considerations, all of which are relevant to the decision. One of the lessons that comes out of looking at this formally is that it's very important that you not omit important considerations from the definition of value criteria.

Many people will ask "why do you say it's scalar?" Okay, the reason you say it's a scalar is that it's the simplest possible way of defining an ordering among alternatives. If you depart from the scalar definition, you can make lots of work for your mathematician but you don't make the decision process any simpler. So, fundamentally, it turns out as a scalar formula.

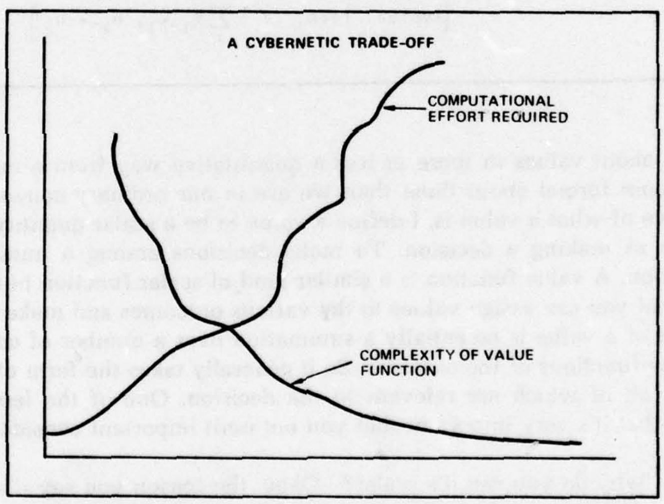
The next point is to try to make clear how the values really are used within a decision process. If we were going to program a decision system on a computer, which would automate the decision process, probably its simplest form would look something like the following: (Slide 14-2) an initial loop which inputs data and fundamentally updates the model or the information about the environment within which decisions are to be made. The next step is simply a matter of considering a series of alternatives, and, for each alternative, in effect, trying to simulate or project the outcome, and, finally, using value criteria to assess the outcome. So, to be able to apply the value criteria you need a way to calculate the value of each of a number of possible outcomes.

One of the questions that arises in projecting outcomes is "what do we define as an outcome?" If we try to think too far ahead, the process becomes very complex. So almost all practical military decisions are made in terms of outcomes that are projected for only a relatively short time ahead.

SLIDE 14-2 THE HUMAN DECISION PROCESS (A standardized or canonical taxonomy).



SLIDE 14-3 HOW FAR TO THINK AHEAD

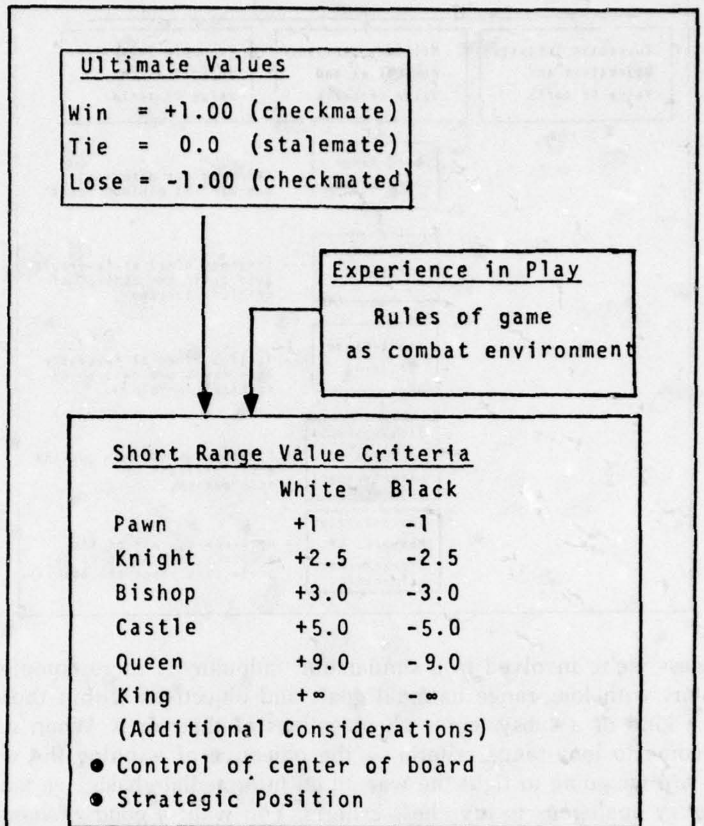


I'd like to think of the problem a little bit like this: Typically, if we try to make decisions by projecting all the way to an ultimate outcome, we're at far end of the scale in Slide 14-3 and we're thinking a long way ahead. The definition of value criteria is fairly simple, you know. We know what the objective is, but the computation necessary to decide what is the best course of action is immense. Now, at the other extreme, you have what people think of typically as rule-of-thumb criteria. You don't think ahead at all. But, given a particular state of the environment, you make some specific decision, and, to make good decisions that way, you end up with a terribly complex network of decision rules, or, alternatively, a very complex value function. Most of the practical decisions that we make every day are made by thinking a little way ahead and using judgmental value criteria to evaluate the projected outcomes.

Now, I want to discuss the relationship between the judgmental criteria that we use to make decisions and the ultimate objective that we may have. The best way to illustrate that relationship is by an example from the game of chess. The game of chess is quite simple in terms of the ultimate objective, or values. The objective is to win, to avoid losing, and, if you can't win, at least to achieve a stalemate. From that very sim-

ple definition of the objective, and from experience in the play and experience in the rules of the game, people generate a set of judgmental value criteria, which are the values of the pieces in the game — a pawn is worth one, a queen has a very high value, a king has an essentially infinite value (Slide 14-4).

SLIDE 14-4 VALUES IN THE GAME OF CHESS

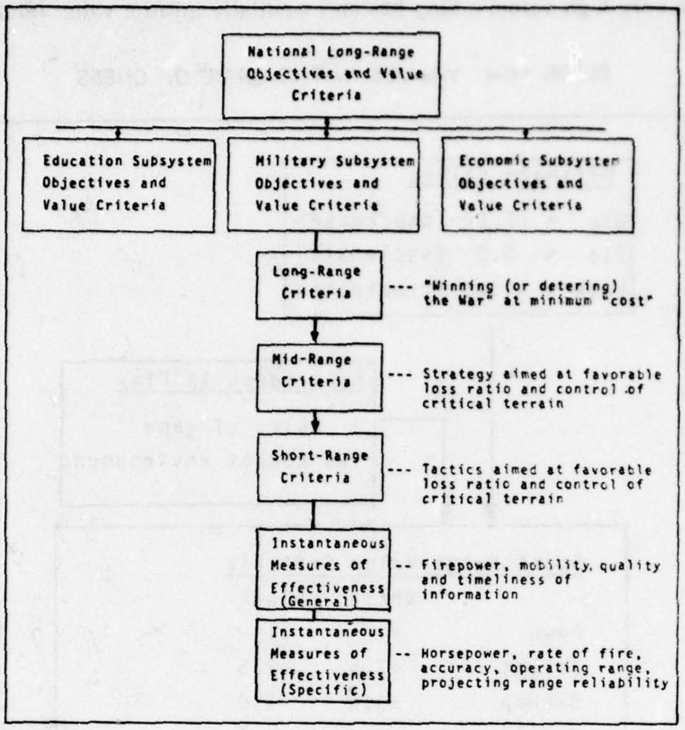


Now, it turns out that, in effect, these value criteria are measures of the combat potential of the pieces on the chessboard. They're almost the equivalent of a firepower score in measuring the effectiveness of the pieces. In addition to these criteria, which are used to make decisions, you try to make moves that result in a good trade-off when values are measured on this kind of scale. There are also criteria that have to do with things like strategic position — control of the center of the board. This is analogous to tactical objectives; for example, "Capture such and such a hill because it's going to get us into a good position for later activities." The experienced chess player makes his decision in terms of a series of criteria, part of which are these simple sort of firepower scores, or the value of the various pieces, and part of which are in terms of the capturing and holding of critical terrain, and so forth. I think that there's a very close analogy between these kinds of criteria and the criteria that combat commanders use to make short-range decisions in a combat situation.

Now, the next question is, "how do these values get generated?" Well, the first interesting point is that there is no theoretical mechanism by which they can be generated rigorously. The game of chess is a particularly good example because it obviously can't be generated rigorously. Formal game theory tells us that every configuration of the board, if one were to play optimally, is either a win, lose, or draw position. So any move you make is going to take you from a win state to a lose state, or, to a transition. These state changes do not have anything to do with these values. These values are useful solely because they enable us to solve a problem by looking only a short distance ahead when we lack the ability to look all the way to the end of the game.

Then, how are the values really generated? They're generated by experience in the play of the game, and by observing the effectiveness of the various pieces in the play of the game. A very similar process is involved in developing information on the combat effectiveness of the forces.

SLIDE 14-5 HIERARCHY OF MILITARY VALUE CRITERIA



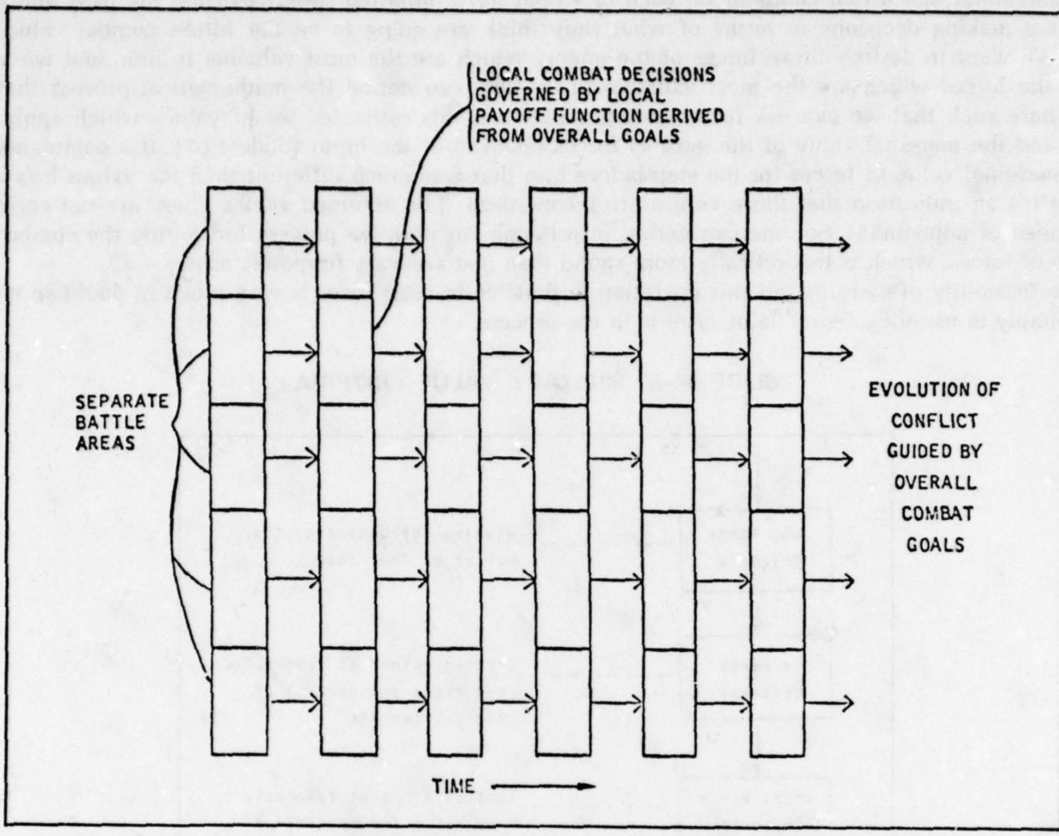
In the military process we're involved in a similar but undoubtedly more complex value deduction process (Slide 14-5). We start with long-range national goals and objectives within those goals and objectives, military objectives are a kind of a subsystem with objectives of their own. When we get into the military business, we probably come to long-range criteria — the objective of winning the war. But, as soon as we move from that to how are we going to fight the war on an intermediate basis, we tend to make the decision in terms of something very analogous to my chess criteria. You want a good exchange ratio with the opponent, and you may have specific objectives in mind, for example, if you're Eisenhower, you want to land on the European continent and take over the terrain. As we proceed downward, however, from this level to the level of individual tactics, we begin to become involved in situations where the objective has to do with short-range tactical objectives on particular pieces of terrain. But these are nevertheless objectives that have to do with exchange ratios.

Notice as I'm moving down this chain I'm moving from value criteria that are useful in terms of looking a long distance ahead to judgmental value criteria that are useful in terms of short-range decisions. As I do that, I begin to lose more and more rigor under this procedure, that is, it's much more difficult to be rigorous about the criteria.

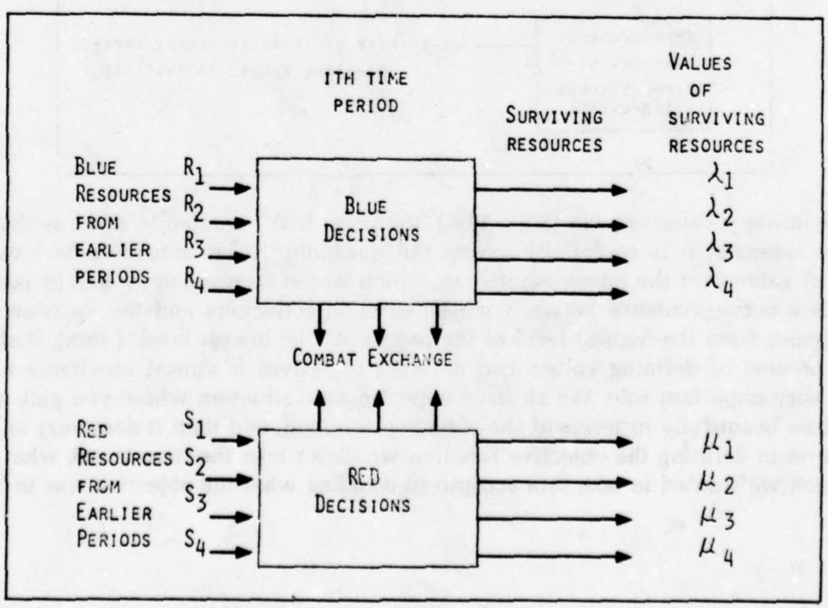
Finally, at the bottom of this chain we come to what I'll call instantaneous measures of effectiveness. These are the measures of effectiveness that tell us what the firepower score is for a tank or what the combat effectiveness of the tank is versus, say, an airplane. That has to be deduced from experience in playing the game, perhaps at the short-range level, perhaps at the mid-range level, and theoretically even at the top level. But, in practice, it isn't practical to do that. So most practical analysis is involved with what amounts to a kind of hierarchy of models. First you use detailed analyses at a relatively small scale to get some of these basic assumptions about the value of forces. Then you play the next level and develop decision criteria at that level in terms of experience that comes at the next higher level. As a practical matter, that tends to be the way we develop our judgmental value criteria.

Now it may be somewhat unsatisfying to look at this process in terms of what appears to be a totally ad hoc heuristic point of view. So what I want to do is put a little theoretical foundation under it. A number of years ago I looked at the process of decision making in a time-sequential combat game. Essentially, I defined a situation where we would have a series of separate battle areas and time moved along as shown in Slide 14-6 and we were interested in making good decisions in each of the subareas in terms of an ultimate objective of winning the war.

SLIDE 14-6 BASIC CONCEPT FOR SEPARATING PROBLEM



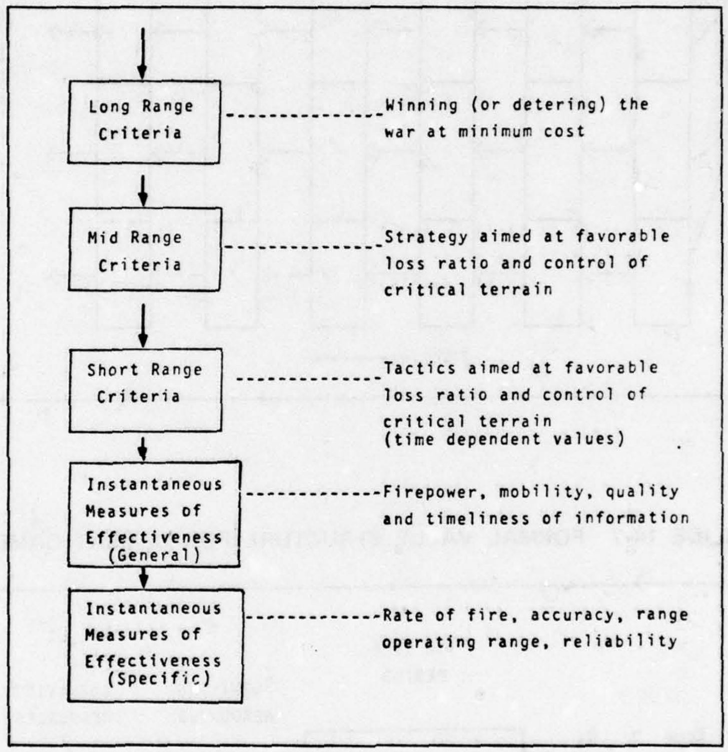
SLIDE 14-7 FORMAL VALUE STRUCTURE FOR COMBAT GAME



When we look within a single cell of the game I just played out, we see that within each cell we have a Blue commander and a Red commander, each of whom have inherited resources from the past, and each of whom are making decisions in terms of what they think are going to be the future combat value of the forces. We want to destroy those forces of the enemy which are the most valuable to him, and we want to protect the forces which are the most valuable to us. One can define the mathematical process that is involved here such that we can ask for consistency between this estimated set of values which apply in the future, and the marginal value of the various forces received at the input (Slide 14-7). If a commander projects a marginal value of forces for the step before him that was much different than the values he's used in the past it's an indication that these values are inconsistent. The assumed values, then, are not correct and are in need of adjustment. So, one can define in principle an iterative process for getting the combat effectiveness of forces, which is theoretically more sound than just arbitrary firepower scores.

The feasibility of carrying out this operation in practice in most cases is very much in doubt so it's often better simply to use judgmental value criteria in the process.

SLIDE 14-8 MILITARY VALUE CRITERIA



Regarding the military value criteria (Slide 14-8), the point that I wanted to make is that, when one talks about an outcome measure, it is essentially asking the question, "what outcomes do I want?" How can I define a measure of value over the possible outcome, which would correspond as well as possible with what I want? So, there is a correspondence between measures of effectiveness and the decision criteria, and that correspondence applies from the highest level of the process to the lowest level. I think it should be apparent by now that the process of defining values and defining objectives is almost inevitably a matter in which judgment plays a very important role. We all have experienced a situation where you pick an objective function, the system does beautifully in terms of the objective function, and then it does very silly things. In most cases, that is because in defining the objective function we didn't take the time to ask what were the various considerations which we wanted to take into account in deciding what the objective was to be.

Panel Discussion — Session II

Dr. Bracken: Let me first introduce Dr. David Dare from the Defense Operations Analysis Establishment (DOAE) in Britain. Davie Dare has been in the analysis business for a long time. He spent a year with OSD Systems Analysis. He's been serving recently as a general purpose problem solver and Assistant to the Director of DOAE and to one of the Deputy Directors. You heard David mention the NATO deployment model earlier. He is one of its coauthors, and I believe he and James were coauthors of the Anti-Potential — Potential method which was used in the IDAGAM and WEI/WUV systems. So he's made many significant contributions to the field.

Dr. Dare: Over the last ten years or so I've been concerned in a number of different ways in analysis of land/air problems related to the Central Region battle, both as a model developer and as a model user in analyses and studies. Today I'd like to say a little bit about the relationship of model building to that of the model user. We've heard quite a lot today about the problems of model building in the sense that it's difficult to represent many of the activities that physically take place on the battlefield. But what we haven't addressed is what impact the model user's requirements have on the construction and structure of models. If I can take this opportunity to say a few words about that, it might help to put a new light on some of the models that we've been hearing about today.

From the model user's point of view we can summarize what happens in many studies I think by saying that at the end of an analysis period the project leader reports to the decision maker, or committee of decision makers, and the kind of report that he makes is essentially to the effect that "I have examined a number of different options that you've asked me to, perhaps I've examined one or two others that you didn't ask me to but that I think you should hear about, and my results are as follows." Essentially, we lay out the advantages and disadvantages of each of the courses of action that we've studied.

Now, one response from the decision maker can be a profuse thank you letter, and when you receive that you should be despondent because it probably means he's filed the report and forgotten it. What is a much more encouraging response is when he comes back with a lot of questions saying, "Why did you get this result?" "I don't believe it." "What is your justification for it?" At this stage, you've got to be able to start explaining some of the background of your analysis to be able to relate the particular factors that are contained that led you to the conclusions that you presented.

Now, here's where the kind of model you're using becomes important. If you've been using one of these global models which represent in a great detail everything that has taken place across the whole battlefield in a single model and that has generated results for you at the theater level, then, essentially, what you have to say to the decision maker is," well, you come to my office for a few days and we'll go through the structure of this model and I'll explain it to you." When you've done all this you say to him, "now, do you understand how I got the results that I presented to you?" He'll probably say, "no, I don't understand a word of it." That is a very serious problem. It's difficult enough for one analyst to understand another analyst's model. The very fact that we've had studies presented today on four-way model comparisons and critiques of various models is, I think, an illustration of the difficulty that we have in appreciating what other people have been doing in their modeling. So, if it's difficult for we analysts to understand what each other is doing, it's doubly difficult for the decision maker, who is not a professional in this field, ever to get much of a grip on what's going on in these models. And, if he doesn't understand them you can hardly expect him to have much confidence in them.

Now, the alternative approach is one that General Welch mentioned this morning but which hasn't been referred to very much since — that is the hierarchical approach to modeling in which we try to break down the modeling system at various command levels. Now, if you have used such an approach then you are in a slightly more advantageous position when the decision maker comes back to you with questions about the conclusions that you've drawn. You can, with some chance of success, explain to him just what went on in the top level model. This model will have rather fewer assumptions and rather simpler and fewer interactions between them so that there is some chance that he'll understand and you'll be able to explain to him what were the important factors that drove your conclusions. The chances are he will disagree with some of those assumptions, and then you'll have to focus on those and explain how you made them. This will drive you to go down to the next level of the detail of the modeling to explain to him how you got that particular set of assumptions that he had objected to. This process goes on until such time as either he's convinced that what you've done is a sensible approach and it all fits together logically, or, he objects violently to one of the assumptions that you've made and insists that you change it. When the latter happens, we get into the sensitivity analysis that was mentioned this morning. But now we're not doing 10^{30}; we're just doing those that are demanded by the decision maker because he's uncertain about some of the assumptions

that you've made.

So, the plea that I'm making is that we try and simplify our models to the extent that it is possible for our customers to understand what goes on in them. Now, I'm not saying that we should never have detailed modeling. But I regard that as a necessary evil that we have to go through to develop more simple analytical models that can be run more quickly and in which the interactions are more transparent. In that respect, I agree entirely with Seth Bonder in that we should strive to draw ourselves away from detailed modeling to develop more analytic representations as our understanding develops. This would allow us to produce theater level models in which we have confidence because they would be based on a long history of detailed studies. But they would also be simple enough for us to be able to explain to our masters how they work.

That is my view of the need for detailed models — a necessary evil, a phase that we must go through, but a phase which we must strive to get away from. I have the feeling that some people feel that the standards by which you should judge a model were the degree to which it represented reality, and the closer one got to representing every individual action on the battlefield the better and more acceptable that model would be. I for one disagree with that philosophy.

Dr. Bracken: Let me now introduce Dr. Reiner Huber. Until about three years ago, Reiner was the chief of about one-fourth of IABG, the think tank in Munich, Germany, that does work for the Ministry of Defense. Currently he's a professor in the National Defense University in Munich and an active consultant to IABG. He also was the chairman of the NATO-SPPOSS conference held in Munich about four years ago.

Dr. Huber: Actually, it is difficult to add a lot more to what David (Dare) has already said on the modeling approaches, because it seems to have covered the problem pretty well. But I would add a few remarks on these two principle approaches that we've been talking about — you termed it the overall, or the global, modeling approach versus the hierarchical approach. It seems to me that most of the models that have been discussed in this session were actually of the global type, at least in the stage they have been discussed today, even though Seth Bonder indicated that they may eventually be more aggregated, in which case I see a hierarchy involvement, so to speak.

One comment that Ed Kerlin made, I believe, was interesting to me because it seems to reveal one of the reasons why this global approach has been taken in the U.S. contrary to that in Germany and also in the United Kingdom which, as far as I know, started using hierarchical models in the very beginning. Ed (Kerlin) said that high resolution models at the theater level were of major interest to the users. Well, from this I would conclude it was because the user said "well, this is what I want." You were being responsive to a user's request. It seems to me that we have the same problem in Germany. Mainly people want high resolution models because they think that high resolution models better reflect reality than higher aggregated models, even though, as Dave (Dare) pointed out, this is a contradiction because, according to the view of the user, the model may be more realistic because it reflects more detail, but it necessarily becomes more complicated so he can't understand the model any more. So, you are more or less forced to find some kind of a compromise, which I, or we, thought, the hierarchical modeling approach really offset. Besides, it's a question of principle. For example, measuring the quality of a model in terms of resolution is sensible because if it weren't sensible you wouldn't have a model — the best model would be reality, so there goes the modeling approach.

Another question that I want to raise came up indirectly — it is the question of, standardizing models, or using standard models. I think Mr. Low treated that in the introductory remarks to this conference, and I'd like to make a statement on that. I do not think that standardizing is a very desirable approach. I think standardizing models would actually kill information because what we would be looking for was the true model. Of course, if you're talking about theater level, there is no way of verifying a theater level model, at least not in terms of the criteria or the results at the theater level. I don't see any possibility of an empirical verification of a theater-level model. So, from this viewpoint, what we actually need are more models which you apply to the same problem so that you expose the differences in those models to get more information so that you can come up with better, or more, information for better decisions. The approach that Dr. Hess indicated, that of using models in parallel, would actually be the approach that was desirable. But, as far as I know, on the theater level, it's only been accomplished once, and this was the Germany/UK/U.S. study. I think studies like this should be done not only to compare models but to support decisions in the sense that you'd like to get a maximum of information with different approaches to one decision.

Mr. Low: Are you saying, in essence, that the multiplicity of models would be a substitute for the empirical approach?

Dr. Huber: Yes, yes, most definitely. Well, that ends my discussion.

Dr. Bracken: I would now like to introduce Lt. General Glenn Kent (Ret.), a person who has had an extremely productive career. He was a valued colleague of ours at the Weapons Systems Evaluation Group, as several people have mentioned already. He set up their force studies and analysis program after a distinguished career in the defense research DDR&E community.

General Kent: As I listened today I sometimes had

the impression that we model builders stand around in the midst of chaos and say "now that I've got that straightened out I'll go on and improve things." Well, while the model builders may say that their models are adequate, it's apparent that a lot of decision makers don't think they're credible. One of the challenges is to gain credibility with decision makers. So, much of what I am going to say is in the vein of perhaps gaining credibility, but still, of course, having adequate models. I don't want for us to perjure ourselves just to gain credibility.

First, I'd like to make a distinction between theater-level war gaming and theater-level computer models. Now, I'm for theater-level war gaming. I'm not so sure that I am wed solely to theater-level computer models. Theater level war gaming allows for further inputs and the computer can assist. As George (Pugh) has said, it helps make clear the effect of different decisions as they go along. I abhor the trend of trying to fix it so that all we have to do to run the model is to figure out where the red button is, press it, and away it goes. We should not try to model decision makers. I think it's a wil-o-the-wisp. Who are you going to model? A particular decision maker? Or, a typical canonical decision maker? If you do that, you've violated one of your precepts of detail. I'm against modeling decision makers. I am for having models that tell the results of alternative decisions. Part of our differences may stem from semantics. As usual, modelers are no better than electronic engineers in the way they talk in their own disciplined conversations. For example, we talk about modeling intelligence — what does that mean? Are you going to model the effect of having better sources of intelligence? Are you going to try and model what intelligence will be available? Those are two very different subjects. I'm for one and against the other.

The best thing that theater level models can do is account for the long-term effect downstream of short-term engagements, and we should make ourselves work hard to do that. Again, along the line of George's (Pugh) theory, how we can be against aggregate input is beyond me. Let me explain. I'm thinking about buying some more defense suppression systems to reduce aircraft attrition of the flight of aircraft going in there to attack the battlefield target. I'm going to buy some EF-111s and decoys and on-board jammers and things that locate those radars and kill them quickly. Now, I know enough about that subject to know that if somebody comes in to see me in connection with some theater level model, and I say, "are you modeling that?" and he says, "yeah, in detail," I'll throw him out. I know he's not, because to model that you have to go to somebody like Calspan who has put in four million dollars to open the door to have a bunch of operators sitting there at the scopes watching what happens as the EF-111 jams and the decoys go in — to see if they're credible and all that. Now, from that, after repeated runs, I can get a sense of what difference it made in the attrition of that flight of strike aircraft going into that area. Then I'll call that an aggregate input, and I'll put it in the theater level model. But I'm not going to allow theater-level modelers to model that interaction because they can't, and, if they try to do it digitally without human operators I know that they're faking it and they've lost credibility right there. So how we can be against aggregate inputs is beyond me. To me, the opposite of aggregate inputs is engineering inputs, or something else. There are two meanings to that word. To me, an aggregate input can still be in detail down to company level, as you described, John (Bode).

*Mr. Bode:*Yes, I think the real key is to get the relevant level of aggregation.

General Kent: I agree. I agree that it's down to company or battalion. Maybe I've got the wrong word for the input I'm talking about, but I want an aggregate input from these levels. So, you can see I'm clearly on the road of a hierarchy of models. I've got to be. I also want to use the theater level model to tell me the different concentrations of troops because, as we think more about theater nuclear weapons, of course, the effectiveness of theater nuclear weapons is directly associated with density — of vehicles, of troops, or whatever. Also, as we try to invent better wide area munitions, the idea of density becomes a very important driver as to their effectiveness. There is a big debate about the effectiveness of forces if they disperse, and I see a great role for theater level models to assess that effectiveness if they're player assisted. I'm all for player-assisted theater level war gaming with great computer assist.

One thing that reinforces my view on that is the exciting thing I've heard lately from General Welch about his TAC EVALUATER. I'm sorry that no one described that today. I don't feel adequate to describe it because all I know is what I heard from him in a short conversation. But, as I gather it, what some enterprising, bless his soul, Air Force guy did was to go out there to Leavenworth and take the printout of a whole stack of documents of player-assisted theater level war games that they had run, like a sand table exercise, and then just put it on a computer so that people could get great visibility as to what was going on. If you want to know what happened back there, you're going to have to thumb through a lot of books. You can see that what he did, first, was the great job of automatic data processing, and from that, though, the Leavenworth people are really beginning to accomplish something. But they got the basic structure of what happened in the war games from players and now, you see, if I change forces too much, why, pretty soon I have to go back and recycle them and have the players say "well, in this circumstance it might be different." But I think that's an exciting avenue to pursue. We've got to figure out some way to at least go up

to decision makers and, for openers, have great credibility on the big things. These things are the high leverage of missions, the high performance of certain missions or functions such as what difference does it make if we double the number of Soviet aircraft we can kill in some raid against our high value targets, what's the effect of that downstream? What difference does it make if, by disruptive attacks against Soviet airfields, we have the number of aircraft sorties the Soviets can generate the first and second day? What is the effect of that downstream against broader measures of merit? What difference does it make if we introduce a fighter aircraft which if when jumped while engaged in ground attack, jettisons its bombs and now has a high probability of killing the enemy interceptor? This, of course, is distinct from an aircraft that does not have good air-to-air capability and can only try to disengage under these circumstances. What difference does it make if you base SAMs on ground launchers or on A-10's (as a splendid first stage) to increase mobility — lateral mobility? What difference does it make if we introduce a better concept for attacking Soviet maneuver units in battlefield interdiction? Of course, also it's obvious what differences alternative force levels make.

While we're on the subject of force levels, to gain credibility let's not ever tell anybody that we can help somebody like a seasoned man in the Pentagon figure out what force levels are required. Commander's appetites are insatiable. If you tell them they can finally do that job with this force, they want double the force because they read about Lanchester and because they know with double the force they reduce the casualties in doing a job. So, there's no end.

What we must be able to tell them to do is to say, with alternative forces, this is the outcome, and that's a different subject. It's the same calculation perhaps, but with a different twist and a different philosophy behind it.

Dr. Bracken: Now I'd like to introduce Mr. Mike Sovereign, who is Chairman of the Department of Operations Research at the Naval Postgraduate School. He worked in OSD Systems Analysis for awhile and worked in the strategic deployment area as I recall.

Prof. Sovereign: It's a tough job to follow General Kent, but I'd like to crawl out on a limb and do a little review of what we've heard today and try to get back to one of the objectives of the conference. I was on the committee, and we did have some objectives that are useful to review here.

We've heard the users telling us one thing but to a large extent the modelers were not quite listening, except possibly for some parts of the last two talks today. We heard General Welch talk about needing real decisions by real people. We heard Phil Louer say that any model to be acceptable to the Army needed to have an identified command organization, which recognized doctrine, constraints on status of the battle like morale, inertia, etc. We heard Bob Schneider's worry about nontransparency and some of the problems that come from it. Andy Marshall, of course, was, as he said, the most negative. He was concerned about the lack of C^3 modeling, for example. Frank Kapper was concerned about the problems of time, cost, and value. And, there were several questions from the audience on things like political delay in a model, or morale, or chemical munitions — things which I think demonstrate that the users want a certain richness in the model and yet something they can understand.

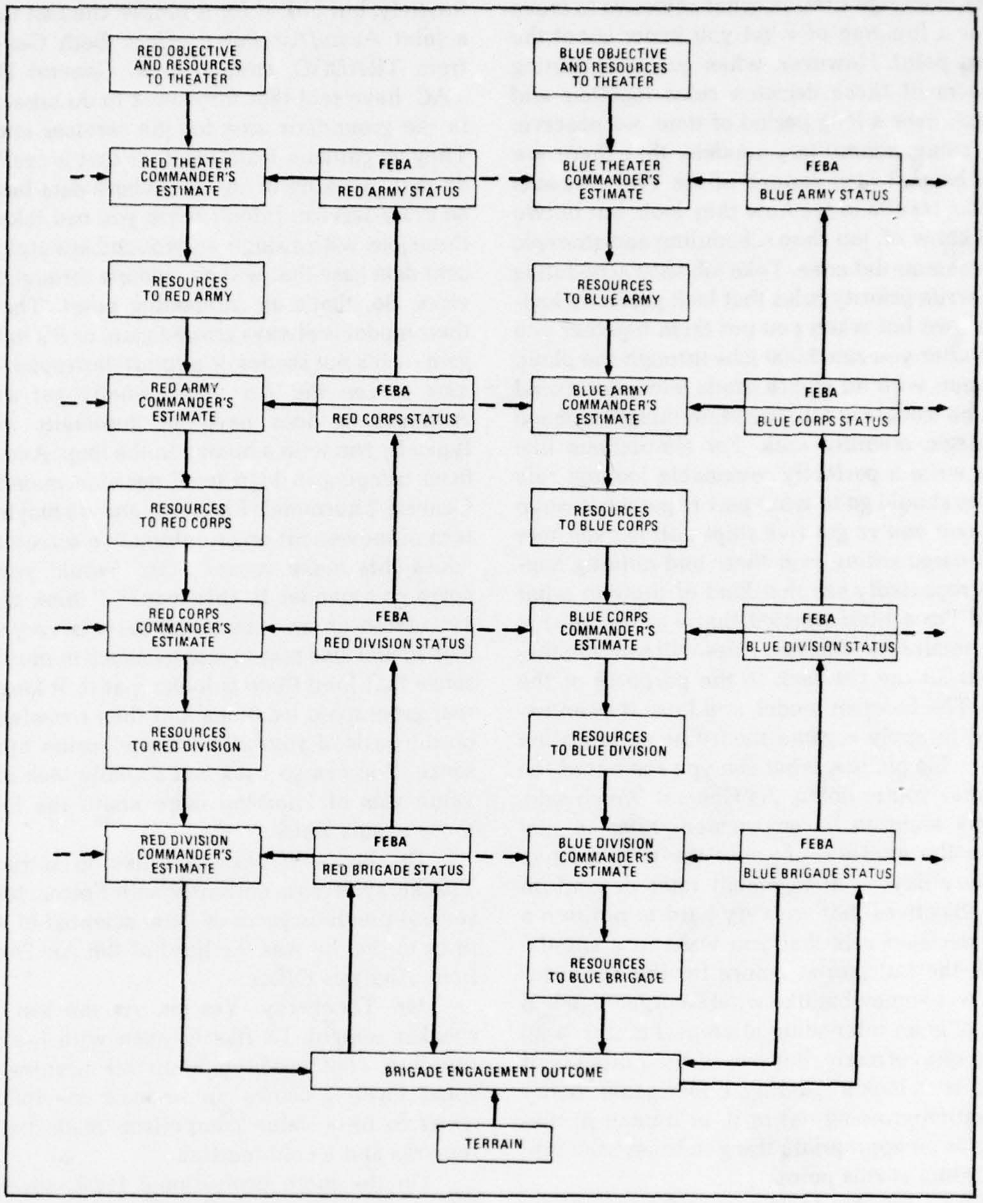
I think probably the most disturbing comment was Phil Louer saying that the major model used is CEM. I suspect most of you know quite a bit about CEM, but to show you what I mean I'd like to use this diagram (Slide B). The only good thing about CEM is the part that reflects the command and control at various organizational levels. The only complex part of CEM is really ATLAS done over many times.

Mr. Louer: I resent that.

Prof. Sovereign: Compared to the sophistication of Vector or IDAGAM, or a John Bode model, you know, CEM just doesn't have it. Yet, the users are using it, and I think the reason is because, although it doesn't do anything fancy in attrition, it does model in a fairly realistic manner the command and control structure. And, most of the decision makers we're talking about — our military men and high level military officers — are interested in command and control. Questions about commitment of forces are what war is about at high levels — not a range to a tank. That's the kind of question they deal with when they're field officers or company officers.

Well, I didn't hear a lot of response to this kind of command and control problem in the modeling that was presented. I think this is a very important area. One reason that I think so is that we have a new curriculum at the school on a joint C^3, and I'm finding out how little there is that is known about that subject. When I was at OSD we had a strategic deployment model. It was a very large-scale LP model; in fact, it was one of the largest at that time. It does everything optimally, which is very nice, and it has a data base that was built up over a very long time with great sweat and lots of cooperation from the services. But it was hard to sell that model. It converted everything to kilotons and moved them. It didn't worry about things like unit integrity — the problem that John (Bode) spoke about: once you aggregate up you can't retrieve how the 82nd Airborne went from A to B. So, we decided we'd take the same data base to get interoperability as General Welch mentioned, and we'd do just a simple loading model. By this means we will have identified that you want the 82nd Airborne to go from point A to point B no later than day 10, if at all possible. By keeping the LP model around you can check that you got reasonable answers. That model

SLIDE B CONAF EVALUATION MODEL OUTLINE.
Note: Dashed lines indicate imperfect intelligence.



gave reasonable answers, and it sold like hotcakes because the military people could look at the output and say, "Hey, that looks like what I do. I can tell them 'go' and I can see it carried out. I know what unit, what aircraft, what ship it went on, and when it got there." That kind of transparency is just awfully nice. I agree with Seth (Bonder) that there are always assumptions in simple things like bean counts, that are probably worse than the assumptions in the model, but at least those assumptions in bean counts are transparent. We hope that the decision maker, has the experience to make up for translating bean counts into results. If we give him a model he can't understand, we can't hope that he can adjust for something that he doesn't see.

So, let me just say a few words about modeling command and control, because it fits nicely with the four models that were discussed by Lanny Walker and Professor Karr. How can you model command and control?

One way is to bring the human into the loop, which at least at some point was pretty much what CEM did (I'm not sure what it does now.).

Another thing you can do is write a decision rule. This is pretty much what the Vector series does. It seems to make a lot of sense to write flexible decision rules, which can be functions of the state variables at a

given time. You can write a perfectly reasonable decision rule about a small unit, whether to engage, which of two tanks to engage first, or, what objective to move onto next as a function of what you know about the game at that point. However, when you start putting large numbers of those decision rules together and running them over a long period of time, we observe, at least in some nonmilitary models, that there are problems. I haven't seen enough of the Vector results or Vector-like results to see how they look, but in two areas that I know of, job shop scheduling and strategic mobility, problems did arise. Take job shop scheduling where you write priority rules that look perfectly logical on their own but when you put them together you get a result after you run 5,000 jobs through the plant, that a foreman with an eighth grade education could better with no trouble at all. The same thing happened in the strategic mobility area. For simulations like SITAP you write a perfectly reasonable looking rule for what ship should go to what port to get what cargo but pretty soon you've got five ships sitting over here and a lot of cargo sitting over there and nothing happening. We repeatedly see that kind of thing in other models, and I'm a little worried that a structure that depends on small scale decision rules will run into that problem. But let me get back to the purposes of the conference. The Lulejian model, and I use it as an example, tried to apply a game theory or optimization structure — a big picture, what can you see out of the whole in what you're doing. As General Welch said, the problems seem to be engagement rates in real modeling, or the question of "does the war last two days or ninety days?" Engagement rates depend on things like objectives that are very hard to put into a small level decision rule that you write in a simulation. I think the Lulejian structure for the optimization, which was somewhat like what George Pugh had on Slide 15-6, is an interesting attempt. I'm not at all sure it came out correctly, but one of the purposes of this meeting is to look at gaming, either game theory and optimization growing out of it, or human-in-the-loop gaming as an appropriate thing to bring back into combat modeling at this point.

Now, it may drop out again as we get more results that can be analytically expressed, but I think one of our major questions is — can we bring in, or, should we be bringing in, more human, or large scale optimization approaches in combat modeling?

I'd like to conclude very quickly with a few remarks about a model I saw just last week at Tactical Fighter Weapons Center, Nellis Air Force Base. It's the one that General Kent mentioned. It's hard to evaluate a model on a quick look, but it has certain very interesting features that I'd like to point out. It's an attempt to marry a model called TAC GLADIATOR, which is an aggregation from the TAC WARRIOR model that General Welch mentioned earlier, at a higher level of aggregation to a ground model which is based on an old CNA model named TWISP. I don't know a lot more about the ground model, unfortunately, but I liked for example, the fact that this was a joint Army/Air Force effort. Both General Starry from TRADOC, and, I think, General Dixon from TAC, have said that they want to do mission analysis in the ground/air area for the services and the DoD. They've gotten a team together that is really trying to do that, working in an area where data bases have to be cross-service. I don't think you can take on one of these jobs with a single service and attempt to get a decent data base that will be credible throughout the services. So, that's an interesting point. The payoff in their model is always ground gain, or it's in the ground gain — it's not sorties or aircraft destroyed — which is nice to see the Air Force finally get away from. Although it does have an automatic mode, they typically run with a human in the loop. And, they have been bringing in high level decision makers, such as General Thurmond, I believe, and displaying the pattern of movement on an interactive screen and saying "does this make sense?" Or, "would you play the corps commander in this way?" I think that kind of evaluation on an interactive basis is very important. The model also plays reconnaisance in much the same sense that John Bode is looking at it. It keeps track of real geographic locations and then revealed locations on the basis of your allocation of sorties to reconnaissance. You can go back and actually look at what the value was of knowing more about the locations of those enemy units.

Dr. Bracken: Now, I'd like to introduce Ross Thackeray. Ross is currently with Ketron, but he spent several previous years as chief scientist of AFSA and prior to that he was the head of the Air Force Operations Analysis Office.

Mr. Thackeray: Yes sir. As the last scheduled speaker tonight, I'd like to open with the subject of humility. That was brought up this morning. On a personal level it comes up because inevitably there's going to be a value comparison made between my remarks and a cold martini.

On the more professional level, several people this morning claimed credit for modeling, for example, by introducing the subject of the need for shelters for aircraft in a NATO environment. I spent some years in the NATO environment, and back in 1956 when Jack Jenkins from the Rand Corporation did a five-foot shelf of books, with no support whatsoever from any kind of a global model, — it was all essentially Friden calculators on NATO desks — with which he proved conclusively that indeed all of the nations should at that point, by almost any measure of effectiveness that you could possibly want including bombs on targets, aircraft loss, and exchange ratios, stop buying airplanes and buy shelters. He was totally ignored.

I submit that the success recently enjoyed by modeling has less to do with the sophistication of our

models and the impression they make on decision makers than the fact that we are suddenly running scared because we no longer have a total nuclear dominance and expect maybe we're going to have to fight a conventional war against at least equal forces. So, there is the decision making context into which each of our studies is necessarily going to go.

Seth Bonder mentioned the requirement for validation of our model results; in fact, he was the only one that did mention it. He suggested perhaps having a contract to look at the Arab/Israeli war, or something of this sort. This is a terribly difficult thing to do. As numbers of people have pointed out in the past, the depth of knowledge you require of the transactions that led to whatever it was that occurred on the field has to be very profound. An unclassified example of that might be the handling of the battle of Lookout Mountain during the Civil War, which violated everything we have ever modelled in the way of the effects of terrain on the FEBA movement, the suppressive effects of fires, and so on. I don't quite know what made all the difference at Lookout Mountain — perhaps it was morale, perhaps it was a patch of fog. But a one shot attempt at validation via combat would be a very difficult thing to do.

There are things which are not written in the histories which sort of throw a little english at least on the conventional wisdom conveyed by the published numbers. But starting with the remarks of (John) Bode, that the types of questions to which we're addressing our attention are changing, I would suggest that there is an awful lot of validation we can undertake without having to go to validating the entire theater model all in one whack. Most of these things, if we take the hierarchical approach, or even if we just look at the component submodeling that goes on inside a global model, are a set of entities whose performance we can validate the modeling for. To some extent we do this. Some people have talked about the use of exercise results. This is one of the primary sources of information. This sort of validation of course, becomes even more possible as we break things down further in a granular battlefield and the aggregations become smaller.

But this brings up the question of who does what in this validation process. It's all very well to cite the results of exercise analysis and say that you got your numbers from that. Well, some of us are still World War II vintage ops analysts at heart and we know very well that, unless you are almost personally involved in the experiments or the exercises, it's awfully hard to extract meaningful aggregate level lessons that will hold up well for large-scale application in a model that may have many users. For this reason, I, too, like (John) Bode's comments about the treatment of C^3, and in fact, here's an area that also fits with what Glenn (Kent) was saying, that we probably don't want to eliminate the presence of the decision maker or try to substitute for him, or model him per se. I think it's reasonable to try to model some aspects at least of the decision making process and of the decision making support facility. These aspects can be an input to the kind of thing that John Bode was doing, or the kinds of things we're attempting to do in the Navy, using something called the Tactical Command Readiness Program, which is a cycle of various graded levels of war games and programmed instruction, to derive the requirements for various kinds of command control support facilities.

I think this kind of cycle interaction back and forth between actual operations, or, as close to actual operations as you can get, and the modeling community is probably very useful. Finally, the last actor in all of this is the model user. So far, we've talked mostly about a circle of people who have had a modeling requirement laid on them, have studied what was available to support the modeling in the way of data, have gone ahead and modeled, and then perhaps done a study, and hopefully the study has influenced some decisions. This is the short term of life of a model. After about two turnovers in a military headquarters, or in an analytic organization, you have people who were not there when the thing was designed, who don't fully appreciate all the short cuts and approximations that were taken, who under pressure of the current crisis to get a study out are quite capable of ramming things into a model that are not totally appropriate to the purpose. I really think we ought to put a feature in most of our models that allows for some kind of self-destruct a few yards down the road.

Dr. Bracken: I'd like to make a few pessimistic comments about what's going on in theater level modeling. It's difficult to do a study that deals with weapon systems of more than one service. At the beginning of this session, I put up a chart on a study we had done with disparate weapons, weapons and replacement pools, Mavericks, shelters, tanks and units with full costs. That sort of thing just doesn't get done very often. Now, there are many reasons for this, and as Bob Schneider said systems analysis is basically a fire drill, I believe, essentially that of answering the mail and crashing on decisions. SAGA has been gutted so they no longer have the strong analytical capability they once had to do cross-service analyses. The Army grinds away with CEM to prepare its long-range plans, but doesn't really do too much in the way of trade-offs of Army and Air Force systems. The Air Force considers its own systems usually, as General Kent mentioned, seldom trading off in Army weapons, and if they do undertake this sort of thing with their Air Force models, it's hard to get the Army to believe them. Likewise, it's hard for the Army to get the Air Force to believe them if they study the trade-offs between the two kinds of weapons. TAC WARRIOR mentioned earlier is a very good model, but it was left undocumented for two or three years while they did

all the studies with it, so it's hard to come in and look at it. They had it contracted out for documentation after a couple of years of use.

Personnel levels are being cut in OSD, DDR&E people, if they have the money, are usually interested in spending it on engineering studies or studies of the specific characteristics of an individual piece of hardware. I really don't see much of a market for studies and analysis using theater level games in the United States to look at serious weapon system decisions, particularly across the services.

But what I would like to see, and I believe the state of the art has arrived at the point where we can do it, is look at the kinds of problems that you were mentioning Glenn (Kent). If you took a couple of the models that we now have and got the same kinds of conclusions independently with an independent data set, that would be a very good argument for taking a certain decision. Seth Bonder gave a talk at Munich about three years ago in which he was saying, "now is the time, we now have the models where we've reached a new level of sophistication," and I believe that. I think we have the models, they're in house, they're working, they're tested, and problems could be looked at with these models. But it's not being done and I see no evidence that it is in work programs, so I'm very pessimistic about the application of models in that way.

Now, from an intellectual viewpoint, I'm very much in favor of looking at command control as the next major thrust. I think that's the right way to go. I think it will be modeled successfully. Whether it will be used or not afterwards, I don't know. Seth (Bonder) is shaking his head. He's complaining that there won't be any data. That, I think, is very relevant. Some big money has to be spent on data if people want to use more detailed models. People won't even use the less detailed models that have reasonably good data for important decisions. So, as a professional in the military operations research, model building, and studies and analysis business, I'm not too optimistic on either score — studies and analysis, or the use of the models once they're built.

Another thing that I'd like to mention that hasn't been mentioned is logistics. When Phil Louer put up Slide 1-1, combat support was mentioned. In a division slice of 48,000 people, 32,000 are not troops, and if you look at the cost structure of the Department of Defense, two-thirds of the cost is in support. We haven't any model, to the best of my knowledge, that reflects the interaction of logistics in combat worth a hoot. We've tried to model these nodes and arcs, but the players can always game many kinds of strikes by sending more from one place to the other. Or, if you represent Europe as an aggregated network you don't really get all the roads and links so that you can't represent it in detail. You can't represent the really basic fundamentals of how logistics interacts in combat, and I don't think any of the trys were successful. There's a lot of money there, but I'm not too optimistic that anyone will do that in the next ten years. I'm much more optimistic about command control.

Now, would anyone like to ask any questions of the people on the panel?

Mr. Louer: Yes, I have a couple of comments to make. I think it's fine to talk about the man-in-the-loop in the models, but I'd like to address a couple of practical problems here. Number one, some of the decision problems that we don't understand very well at all are high level decision problems — those at the corps and the army and the theater level. I think these are the most difficult to model at this time. To do that you have to get people who have experience at this level, and this means getting top level people to perform as this man-in-the-loop. Now, you build a model that has the facility for a man-in-the-loop and you turn it over to an agency to run it — for example, CAA gets it in the Army. There are no top level people at CAA. You'll find lower level people acting as top level people. I agree with the objective, and I think it's the thing we need to do, but I wanted to note that there are practical problems involved that really limit this kind of analysis.

Dr. Bracken: Would someone like to address this? Glenn or Mike?

General Kent: You're right. I agree with you that there may be a practical problem here, but you know I thought about that and you'll notice when I said that you have players involved I didn't say senior commanders. Now, I find that it is far easier for me to grasp that I can have senior commanders teach majors how to do that kind of analysis instead of having senior commanders teach some stupid computer how to do it. You see, there's a certain illogic. We both agree we'd like to have a senior commander do it — he could be retired, or whatever. But, we can't do that; it's not practical. We can't have top level people available all the time.

Now, what's our escape valve? My escape valve is to teach majors, and I contend it is better for me to do that than to go to some man to try to teach a computer how to do that, because somehow or another that major has got better algorithms in his head than that guy can ever put in that computer, and they're easier to change too.

Dr. Bracken: Jim Taylor?

Prof. Taylor: I'd like to take partial exception to what General Kent has said. I think there is, or course, decision making expertise in this area that, at the higher levels, might possibly be in the senior war colleges, like the Army War College and the Air University, and possibly things along this line may bring in that dimension.

Dr. Bonder: Let me suggest we tried that. We wrote to some of these places and suggested such interaction. They said you've got to be kidding.

Prof. Taylor: Without going into detail, Seth, I think it's taken a little more serious than you just portrayed here.

Dr. Bracken: Would anyone else like to make a comment or ask a question?

Dr. Karber: One thing that sort of struck me was the highly one-sided approach to the discussion. One side is talking about theater war against Europe and we had a little bit of discussion about the West Germans and British, yet certainly what's happening regards gaming and modeling among the Soviet military is of interest. Not so long ago, John Erickson pointed out that, as a result of some of the extensive work that he was doing, was getting what he thought was just the cream off the top of what the Soviets were getting in theater level war gaming. He made the statement that the Soviets had taken a dedicated team of some 200 people staffed by general officers, who are going back to Second World War data and empirically analyzing every engagement they had any records on. Then they take and essentially computerize the data base, and then go through and start updating it consistently with the new data inputs that came in. It would be very interesting, to me anyway, to hear anybody's comments on what we have seen the Soviets doing. If we haven't been paying any attention to what they are doing that might be a very interesting place to look. I've seen it in personal experience with logarithms and mathematics. They are really quite vulnerable.

Dr. Bracken: There is an operations professional in the Washington area, Al Rehm, who's now with CIA, who reads Russian. For a long time he attempted to keep up on the Russian operations research literature. Could anyone respond to Phil Karber's question?

Prof. Shubik: I can respond to some extent. In Russia, the connection between war gaming and administrative or business gaming is unbelievably poor, and I'll touch on that in a minute. But in surveying what was going on at Moscow and at Leningrad, the first thing that really surprised me was that they were esentially 10 to 15 years behind us and were bent on making every single mistake that we have ever made. In fact, and this is rather amusing, maybe four years ago, in France, a couple of young Russians from the Soviet's main computer center came up and with great enthusiasm described their current plan to build COW. Now, there should be at least somebody in this room who remembers COW. It was built at Rand at about 1956.

One thing that may help the Russians considerably is that their computer technology seems to be still bad enough that they won't make the great error of putting every model in the center of a computer so that it will be untouched by human brains. So, at least on that level, the Russians still have a hope. They cannot totally simulate without human input.

Now, the last point that I want to make is one which I found really fascinating. I was in the Soviet Union exclusively to survey business gaming, administrative gaming, and gaming for teaching purposes, which were quite clearly not military. But the thing that fascinated me was that their security is so good that I am really willing to believe that there was virtually no interception between the people in those areas of gaming and the people doing military gaming. In fact, there seemed to be three different groups. There are the business gamers, with their French academic style. There are the mathematicians, who sit quietly doing differential games, which they do rather well by the way. Their game theory is unbelievably good as are their differential games and their knowledge of the literature. But they sit aside in a completely different area. Then there are the military gamers, who must be somewhere else. I couldn't find any trace of military gaming, except what I would call Mickey Mouse applications by the mathematicians.

Dr. Bracken: Let me ask Bill Lucas, who's spent a lot of time in Leningrad at the Game Theory Institute. Do you have any knowledge about the Russian military gaming?

Prof. Lucas: Well, I would agree about the business gaming. On the other hand, the previous gentleman (Martin Shubik) expressed the fact that there are mathematical aspects to support some of the things the Soviets are doing, and there is some evidence in my mind that there is some strong mathematics supporting some of the military aspects, some of it in the gaming area. Their differential games, their control theory does have a military outlook. To me, it's obvious that it's highly classified, but you can get enough of it to indicate that some of the real giants in mathematics over there have spent some time on gaming and game theory for military purposes. But the gaming that is published is really very disappointing. I mean, they do have a copy-cat tendency in both theory and other things. Now, in theory, sometimes it's good to go through the basics and then back up. But in gaming I agree that they're back 15 years.

Dr. Bracken: Jim Taylor?

Prof. Taylor: To comment on Phil Karber's point I have seen one translation of a Soviet military book. It's somewhat dated, but, even though it's a very small data base, at least it's some data base. I did take time to look through this book, and it was in some aspects a reminder of the late 1950s in its treatment of hit probabilities and it's treatment of optimization. The latter really wasn't plugged into operational modeling of military systems. It's treatment of combat processes was circa 1954. So there we have one translation of something unclassified, which was extremely dated. On the other hand, you do have to realize that their copy-cat editions do get involved in things of defense. For example, Kolmogorov did write the theory of firing after World War II, which was translated by Rand.

Dr. Bracken: Could you recall the citation of the

book you're describing?

Prof. Taylor: It was something that was translated by NITS, it's on the PG school shelf. I don't remember the name, but by random walk I can probably run across it again, or maybe Reiner knows it.

Dr. Huber: Very early German translation shows two books, as far as I know, on Soviet military operation research. I'm quite sure that the books should be looked at carefully to judge their state of the art because my impression is that they were published to show military officers how to use operations research methods in their day-to-day decision making. One book is by Wentzel on military operations research, and the other one is by Chuyev, also on principles of military operations research. Both authors state explicitly that they write their books for military academies, so these are probably sophomore or junior courses on military operations research.

Dr. Bracken: Seth Bonder?

Dr. Bonder: It is clear that there is a dichotomy in the very simplistic kinds of books that come out of the Soviet Union and the simplistic kinds of things they do, which are well publicized, and in the fact that it would appear that the more sophisticated kinds of things are hard to get your hands on. They are done at particular institutes in the Soviet Union that are not open. There are really two levels.

Dr. Bracken: Frank Kapper?

Dr. Kapper: There was a symposium over at State, and I would suggest that if anybody is really interested in this they talk with Al Rehm who was a key figure in the symposium (he was with Ketron then, and is now with CIA). I was a JCS attendee at that meeting, and the extensive literature that he (Al Rehm) had was very impressive.

A particular point that I'd like to make has to do with tempering Jerry Bracken's pessimism a little bit, not too much, but a little bit. In conversation with Seth (Bonder) I was saying that I think, in the modeling community, we have not really gotten our act together so to speak. Right now, I think the budget-cutting in Congress is really hurting us. I think the overall attitude towards studies and analysis in Congress has really hurt us. I think the Dep Sec Def study performed under Secretary Clements has really hurt us. Right now, we are at the low point, if you will. I do feel that this place may be long term.

Dr. Bracken: I forgot to mention my feeling about Congress. Congressmen and the services working together have incentives to keep the OSD capability and any integrated capability to a minimum.

Dr. Kapper: That's the function of the CBO's budget committee, to take the impact of Public Law 90-244. But the point I was going to try to make is that we'll not press on and improve the standards of studies and analysis along with gaming unless we do a better job preparing ourselves when we are called to testify in Congress, particularly in this budgetary area.

Dr. Karber: In addition, I just want to follow up my comment regarding Soviet trends. I'm not sure, and I certainly wouldn't infer, that they're farther ahead or necessarily farther behind us. In reading Soviet military literature in one particular area regarding tanks and anti-tank discussions, it was very clear that there was a method behind their madness. They're using some models, they are setting norms for force requirements and for data to be used in simulations. They're using a simulation technique, going from theory to field testing, and then using those field test results to update the norms. Al (Rehm) suggested that their process could be very fruitful to study. Having just testified on the Hill a couple of weeks ago, and this is really more to the pessimistic comments that were related to the standards of theater war gaming in the United States, the pessimistic potential, if you will, it seems to me that there is nothing easier to sell, whether it's in the Pentagon or on the Hill, than the relative question of ability. I would suggest that it might be kind of interesting for somebody to argue that this might be a worthwhile area to turn some of our own gaming research efforts on — that is, really try to break their operational code. How do they *really* see things? I think that could lead to a lot of insights into the way we arrive at our own conclusions, and support John's (Bode) point about keying on vulnerabilities and the interforce decisions as well as individual service weapons procurement. Also, if we went into assessment with Andy Marshall, I dare say that Andy would be one of the most enthusiastic supporters of a theater level war game development that took as its premise a lot of hard data that is available and is apparently classified, and that essentially tried to "reverse analyze" the way the Soviets were looking at theater war gaming. I think of it as a marketing perspective as we try to keep the discipline alive.

If we assume that our force structures and weapons modernization are significantly impacted by analytical techniques incorporating simulation, and the evidence seems to indicate that the Soviets do likewise, can we then really afford to lose the simulation race?

Dr. Bracken: Jim Taylor?

Prof. Taylor: Possibly a lot of the problems stem not so much from modeling as from the management of modeling and the organizational framework for using and facilitating the use of models. It is important to minimize barriers and hurdles and not to do clever things like A builds the model, B maintains the model, and C uses the models. That's very clever for making sure that you get minimal communication and that resources are used to a small fraction of capability. I've got documented cases that say that's the manager's philosophy.

One other thing, in validation, there are a number of people here who have spent large portions of their life looking at this area. It's a very tough area, I think

they'll agree. I don't think we should confuse real live combat with exercise data. How many people here have ever seen a field experiment analyzed or played, or planned one? I think there are some real problems with confusing exercise data with real combat data.

Dr. Bracken: IDA is very much involved in helping in exercises and there's a great deal of work in test and evaluation

General Kent supported General Starbird with a lot of activity. Of course, the effort in test and evaluation is not closely connected with war games because the war games usually use what should be combat data or historical combat data.

Prof. Taylor: Yes, and a lot of OT&E generates more engineering or performance data than data that are really of use in combat modeling. All I'm really familiar with is Army field experimentation, which is all the way out on the West Coast.

Dr. Bracken: Would any other panel member like to make any comments? Well, then, thank you gentlemen.

Session III — Theater-Level Combat Modeling Methodology

15 — Attrition of Ground Systems Modeling Overview

MR. ROBERT FARRELL
Vector Research, Inc.

Factors of commonality among ground combat models — plus explicit modeling as opposed to implicit treatment of combat effects

Mr. Farrell: I'm really just standing in as chairman for Wilbur Payne, who was scheduled for this role originally and who has since found that he had to be in Europe this week. So I'm really going to be a time keeper for everybody along with giving the paper that I was originally scheduled to give.

In my paper, I'm going to talk about what was originally called attrition of ground systems modeling overview, and I've taken it to be a slightly more general topic, ground combat modeling overview. It's hard to say where the line ends when you talk about attrition methodology — what is included and what is excluded. So, I'm going to be a little more general than some of you might usually expect in dealing with attrition.

I started out trying to take existing theater level models that I knew about — American models. I'm not going to say anything about British or German models in this area because I don't really know that much about their structure. But during the panel discussion at the end of this session maybe we can get some comments on how foreign models compare in ground combat modeling and in some of the other areas that we talked about this morning.

So, for American models, I tried to isolate what I could say in general about ground combat modeling, and I came up with one major conclusion: there's almost nothing in common between the five (six if you count ATLAS) major operating theater level models, in this country.

The one area in which there is some commonality is the concept of how to represent combat forces. All of the models agree that there is a vector of force strengths, the number of tanks still surviving, the number of people, the number of APC's etc. After that, you get some differences about the number of what. All but one model agrees that there is a vector, the number of different kinds of weapons that are still surviving,

which are important. The exception is ATLAS which has no real treatment of attrition of weapon systems, only of people. Surviving people are used as a proxy for all the other surviving equipment, as I understand it.

The organization of ground forces in these models varies in richness, but it is overall similar in concept for the various models. They all treat the organizational hierarchy of ground forces pretty much in the same way, parallel to the actual structure that you see out there in military reality. The type of missions that they treat are an abbreviated set of those that the military recognize, so they talk about delays, attacks, withdrawals and up to maybe six or seven postures. They don't agree on much of anything beyond there. In fact, many of the other processes that relate to ground combat — target acquisition, maintenance, combat engineering — are not explicit. As somebody said yesterday when we were discussing TAC WAR, "well, we really have target acquisition in IDAGAM, it's in the data base." That is common to essentially all of the models except VECTOR 2. There is target acquisition in those models in the data base. It is implicit as far as the model structure itself goes. Somebody has built assumptions about target acquisition, target selection, logistics effects, into the attrition calculations by changing the input numbers. I'm going to talk a little bit later about why I think people have taken that route, and what I think the differences are between that and alternative routes in which certain processes are explicit, but let me say there is *that* much commonality at least.

Many users don't, in fact, think hard about target acquisition, target selection and so forth, in generating their data. So we find that these things are implicit and that the assumptions are not always very clear. Regarding one of the comments made yesterday — that it is very hard sometimes to find out what the assumptions in an actual study were. I think part of the reason is that they're hidden in the data base. Talking about data bases, the data used in these models is non-uniform in format, similar in ostensible definition, but actually non-uniform when you get into the details. Many of the models have things that would be called attrition coefficients, or attrition rates. They aren't the same attrition rates in the different models, and you can't take a number that goes into one model and put it in another model and expect the model to make sense.

The attrition equations proper are different in every model — no commonality at all. I can't speak for IDAGAM I and IDAGAM II, I don't really know much about those, but even between VECTOR 1 and VECTOR 2 there is commonality in some areas but not in all. When you move from VECTOR 1 to VECTOR 2 the attrition equations change in some areas. I think we learned some things in between, but it makes it very hard to say anything in general.

In some recent work for the Army, I did have an opportunity to learn that, in spite of all this lack of commonality in ground combat attrition, there are some things that have happened that I think we could learn something from, and that there is much more commonality really there than one would think.

So I can talk a little about that. I'm first going to try to define some terms. I think we're all familiar with the notion of an exchange ratio, which is a ratio of losses of two different system types, typically on two different sides. We talk about an exchange ratio as the ratio of losses of red tanks and blue tanks. You could also talk about an exchange ratio that is the ratio of losses of red tanks and red APCs, and there might be a characteristic ratio. We'll see in a minute why I think that might be true.

Slide 15-1 — DEFINITION OF TERMS

Exchange ratio

$$\frac{\text{losses of system 1}}{\text{losses of system 2}}$$

Conditional exchange ratio

$$\frac{\text{losses of system 1 in condition C}}{\text{losses of system 2 in condition C}}$$

Conditional exchange direction

$$\frac{\text{losses of system i in condition C}}{\text{losses of system j in condition C}}$$

These are typically, originally at least, overall concepts (Slide 15-1). You want to talk about an exchange ratio between the global measures of the strength of two different sides. It can be more specialized, and I call

the specialized version a conditional exchange ratio. This is an exchange ratio limited to a particular kind of engagement under special conditions. On the slide I call it condition C. This may be delay engagements with force ratios between 2:1 and 3:1, or it might be air-to-air combat with three penetrating aircraft and two intercepters, or any other nominal set of conditions under which we might expect similar results, and we could talk about the exchange ratios characteristic of those conditions.

We still have talked only about two systems, but in many attrition exchanges in these models, and, I certainly think in actual combat, we have many systems out there interacting at once. This is particularly true in the ground area. For these interactions, I have what I call a "conditional exchange direction," which is the set of all these cross proportions. Instead of saying the exchange ratio between red and blue is 2:1, one can say red is losing two tanks for every tank that blue loses. He's losing 0.4 BMPs for every tank that blue loses. Blue is losing 0.7 APCs for every tank that blue loses, just to keep the denominator constant for a moment. One can have a set of characteristic ratios of this kind, and one might think, as I have thought for several years, that it would be very nice at least if those exchange directions really were fairly constant over some reasonable class of conditions. That is, if one could say delays had a characteristic kind of exchange.

Well, working with some people in DARCOM, doing some analyses based on games that had already been run, I had occasion, to see the results of five different division and higher level models on Central Front campaigns. We were looking only at one day's results, so what I'm about to say is somewhat limited. I don't really know very much about how these models would talk about downstream effects, but in similar Central Front campaigns, and, when I say similar, each of these models had been run on its own data base, developed independently for a different study. These, then, are five different studies from which we have taken results completely independently. We looked at the forces they had and they all described basically the same war, not very surprisingly. The time frames were a little different, the assumptions about the exact weapon systems present were a little different, but basically they all had the same force ratios and very similar weapon mixes. They gave what I call general agreement, within about plus or minus 10% on these conditional exchange directions (Slide 15-1).

Now, this was true not only of ground but of ground and air, for those that model air. JIFFY game results, and maybe DBM I'd have to go back into my notes — did not include air.

In spite of the fact the five models all have five different sets of equations, and that I'm willing, as are a lot of the rest of us, to go out and spend an afternoon with somebody arguing about which set of equations is better, we all seem to have a set of equations that does describe something that we really agree on. We heard something like this about the tri-national comparisons yesterday, if I understood properly what I heard. I think Dr. Goad from STC was saying that they had seen a result similar to this in comparing one American model — I don't know which one it was but I'm assuming it was IDAGAM — with a British model and a German model.

We didn't get quite the same result that I think the tri-national comparison people got. If you talk about conditional exchange directions or ratios here, clearly the remaining question is, "well, what about the rates?" I think what they said in the tri-national comparisons was that the battle was just faster or slower in the different models, and they included comparisons of FEBA movement with attrition as well. So, they were talking about something slightly more general than the exchange direction that I'm talking about, which is just attrition. What I had a chance to look at was just single day results. Some of these studies used periods longer than a single day, but not all of them, and we did not do any comparisons or analysis on the longer periods.

We didn't find that they agreed with rates, so we had the same result the tri-national comparisons had to that extent. They didn't agree in overall rate. Let me see if I can get that Slide (15-2) up where people can see

Slide 15-2 — COMPARISON SUMMARY

Similar central front campaigns examined using:

CEM, IDAGAM, VECTOR-1, JIFFYGAME, and DBM

gave general agreement on conditional exchange directions. (±about 10%)

They did not agree with respect to overall rates of activity or relative rates of different activity.

The largest difference was that IDAGAM differed greatly from all four other models.

it better. The overall rate of activity was different, but further, and this I think is a different finding than the tri-national comparisons, the relative rates of activity in the games were also different. Not only were some games slower than others and some games faster, there were games in which air activities were fast and

ground activities were slow, and others where ground activities were fast and air activities were slow, and not just air and ground. Within the ground itself, the relation between artillery-related attrition and direct-fire attrition was different in different games.

We did do some analyses on the overall rates and four models agreed within about plus or minus 15% on overall rates. They also agreed much better on relative rates although there were significant differences, much larger than 15% in some cases, in comparing the rates of some of the artillery-driven activities. We believe that was due to differences in assumptions about whether or not Red had ICMs. We're not sure; it looks very like it from some of the things we saw, but we did not have complete descriptions of the data bases for all five of these cases — only statements that they were developed from genuine intelligence data.

The one outstanding model was IDAGAM. IDAGAM differed significantly. I talked very briefly with Bruce (Anderson) yesterday evening about some of this. Let me say, at this juncture, that this project was not a methodology project. These observations are side observations. They're off the topic of what we were actually doing in this study, which was to try to solve an analysis problem. In fact, all five models agreed about the solution of the analysis problem as we saw it. When asked which weapon system changes of those we were looking at were, in fact, more important than others, they all agreed, which is another nice feature of this comparison.

But let we say again, IDAGAM differed greatly. The principal differences were in the comparison of really three sets of rates: one was the direct fire exchange of ground systems with respect to the attrition of weapon systems; one appeared to be a rate having to do with personnel attrition on the ground; and, one appeared to have to do with air and air defense systems. Now IDAGAM appeared to have clocks, if you want to call them clocks — rates at least — for those three activities which did not correspond. Specifically, the ground weapon system was quite a bit slower than any of the other models. The ground personnel attrition was somewhat faster compared to weapon system attrition. I don't know why, it may or may not have had anything to do with the models. Each of them had its own data base. It may have been in the data bases. In any case we had very good agreement among all the models on conditional exchange directions. And, I'm going to continue to say that at least, regarding that particular element of combat, we all seem to have a common understanding, no matter how much difference of opinion we have about the equations we use to represent it. And I say that from limited evidence. We have two cases, the tri-national comparison and this comparison that I know of at the moment, but we have no other attempts to compare. At least we don't have any failures yet. So, I hope we have a common understanding of it. The fact that our understanding is common, even though the methodologies we have used to arrive at that understanding are very different, may mean we're right. Or, it may not. We may have collectively led ourselves astray somewhere.

I don't think we're in that position yet on rates. There's more difference in rates even among models which seem to agree somewhat, and there are some very significant differences. I gather they would be even more significant if we tried to put this set together with the tri-national comparisons since, if I understand correctly, the comparative rates of ground systems attrition, personnel attrition, and air-to-air defense attrition, were roughly similar in the tri-national comparisons. Consequently, IDAGAM would not stand by itself, but with two other models in what would be a total set of seven models. We would have a group of four and a group of three, which had somewhat more consistent pictures within groups than across groups, and significant deviations both within and across groups on the questions of rates.

Dr. Bracken: May I ask you to clarify? When you talk about the relative rate of different activity, what is the definition of that, and particularly with respect to your definition of conditional exchange direction?

Mr. Farrell: Okay, we arbitrarily lumped results — taking killer/victim scoreboard results for these. We considered ground direct-fire-caused attrition on weapon systems and said that was a clear lump — that it was a condition, that it was a type of engagement. Air/air defense engagements were lumped likewise. There was almost no ground versus air defense attrition; it was separately lumped as another one, but we didn't really look at it. Then there remained the artillery battery/counter battery duel and the artillery versus personnel section. We took that as a single condition, a type of engagement even though there are really four attrition rates involved — artillery on artillery, two-sided; artillery on personnel, two-sided. So, we have three sets of conditions that we really looked at. The air/air defense engagement, the artillery-related engagements, and the ground direct fire engagement.

Dr. Bracken: Within and between those engagements what about agreement?

Mr. Farrell: Within those engagements we had good agreement as to the directions of exchange, the ratios of exchanges. Furthermore, we had reasonable agreement, although somewhat broad but still plausible agreement among rates with four of the models. But not with IDAGAM; across those rates we got much less agreement anywhere. But, again, IDAGAM was sort of the outstanding example because of the rate differences both overall and in the pattern of the three rates compared to one another. I have no idea, you

know, whether this observation really means all that I am interpreting out of it. It is a very limited sample of evidence. I would like to see more people do more comparisons of this kind and see whether we do have an understanding of war. I think that really is one of the things that if we were to be able to realise that understanding that it would lead us finally to a simple transparent model of some of these praxes. If, of course, we can learn what our more complicated models are in fact predicting.

Dr. Goad: Bob, I think I'd perhaps better amplify what I was saying yesterday about the similarities between the FRG, UK, and U.S. models. What I call the rate of progress of battle is different from what you're calling rates of activity. Without going into the detail of what the overall results were, the three models that I looked at had many activities within which, indeed, different rates were incorporated for things like close air support, the numbers of sorties that were flown each day, and the number of forces that were committed at any particular part of the Central Front on any particular day. The name of the game was to try and normalize out those differences and then ask oneself the question "What would happen if the three models, different though they were, had been applied to the same battlefield situation with exactly the same conditions? What would have been the rate of progress of the battle?" That's what I tried to do, and in spite of all the differences within the individual force connections, I came to the conclusion that indeed the rate of progress was the same. But, when you are looking at five different bases, there is a big problem in trying to normalize things so that they are, in fact, comparable.

Mr. Farrell: That's right, although my problem is less than yours because I am looking only at one day and not the long-term rate of progress. Let me see if, in fact, I'm correct; you did see that what I'm calling *conditional exchange directions,* the ratios of exchanges given comparable conditions, were essentially the same?

Dr. Goad: That's true.

Mr. Farrell: Okay, so we do see common agreement at least on that first one. I may have been overinterpreting some of what I thought I had heard on the rates, so it might still be that your results are consistent with ours on some of the rate questions.

Dr. Goad: Well, when I talk about rates, I'm not talking about how long it takes from start to finish. I define how long it takes for comparable events on the battlefield to occur; in other words, how long it takes for the WARSAW Pact to achieve some human objectives, never mind the forces that are committed.

If we normalize out differences so that everything is constant regarding the data inputs, then what would the models say? That's the question I was trying to answer. I came to the conclusion that the main difference was simply in time.

Mr. Farrell: Okay, I think that confirms some of what I'm saying here, and it may or may not differ with my final point about the relative rate of different kinds of activity. It sounds like you didn't really look at that question — you tried to normalize out whether the air war happens faster than the ground war in some models and not in others. You essentially said, "Well, let's try to stabilize the conditions we're looking at." Okay, so we may have a set of essentially consistent results. We sound like we certainly do on the first point, which is the one on which we have the most coherence anyway. So, I still feel that we may be coming to some kind of common understanding that we still haven't codified in a form that's easy to use — possibly, we haven't reduced things. We have the sets of equations from each model, and we now know they behave similarly, particularly in the games we really are concerned with. We haven't really sat down and said what is a simple model of that behavior that they all could agree with and that summarizes it without the levels of complexity that we see is in the four-model comparison, for example. That's the main point I really wanted to talk about.

But, also, I want to talk a little bit about implicit treatment versus explicit treatment since in the ground combat area there appears to be agreement about implicit versus explicit. The agreement seems to have been to go to implicit treatment of effects in the data base in many of these models. I think there is a reason behind that — what people have called transparency.

Here, I'd like to introduce three terms. The first is the visibility of model logic. What I mean, and what I think most people mean, by visibility of assumptions is that there is explicit acknowledgment in the model logic of some particular feature. So, when someone says "is the target acquisition process visible?" that means "Do I have an explicit statement?" Perhaps it's in a preprocessor, perhaps it's in a statement about how to generate the data, but is there at least an explicit statement connected with the model about what the effects of changes in target acquisition function performance would be? That's what I would say means visibility.

Now, people have also talked about a concept called *transparency of model logic.* I would argue that what they mean is that they can see and understand this model, if not at a glance, at least with a limited amount of study, and I would argue that visibility and transparency are in direct contrast with one another.

I would say, in fact, that visibility of all the parts of war produces clutter. That is my third concept. When you try to look at the model instead of transparency, you have too many pieces in front of you. I don't think I'm extending the metaphor too far to say that really is what happens to me when I look at these models — that we have what are really conflicting pressures in trying to make sure that all these parts are in fact visible, that our assumptions are very clear, that the structure contains all the things we know will have credible effects, which we can vary for future users of the model who are going to come in with a problem we don't yet know about, that we can treat the particular problems of future users and their particular processes with explicit visible items, and at the same time, that they can see right through them. So, the reason why I think we have a real conflict in implicit versus explicit effects is because implicit effects (1) do not clutter the model's transparency, (2) do not allow for dynamic variations in the effects with game results, and (3) require understanding of the model in input preparation.

Implicit effects don't clutter the transparency. If I put the effect in the data base then anybody who asks me what the model does doesn't have to hear about the effect. However, these are major problems. I haven't talked about the advantages of being explicit. Let's talk a little about what implicit effects don't do for you. They don't allow any changes in those effects as the game goes on. So if I build target acqusition into the data base my target acquisition is the same on day one, day two, day ten, and day thirty, in spite of the fact that I lost 30% of my force in one game and only 10% in another. Target acqusition stays the same even though I get information during the game. This remains the same in the long run, despite the fact that I actually learned something about the other side's behavior and can do better because I understand it better.

If I put command control target selection into the data base, I am again assuming it goes on the same in spite of the fact that I've killed, say, 30% of the other side's tanks. I will select as many tank targets as I did on day one, so there are detriments to the implicit effects and strong points that are pushing for explicitness.

Also, as I think I mentioned briefly earlier, somebody has to actually spend the time thinking about those implicit effects in generating the data. I know I have been remiss in what I really would say is the documentation step of models I have been involved with. I include a brief mention with each data item that says you should consider all these effects which we have put implicitly into data base, and then I stop. I think that is true of a lot of other model developers and a lot of other model documentation, and that doesn't really help a user other than the original creator. I think original creators can do a lot of very nice analyses with their models. I think it is very difficult for other people often to do that kind of work and that is part of the tremendous load that people associate with building the data bases originally. It isn't really putting those numbers into the machine that takes three men six months. It's trying to figure out what number it is that really is supposed to be going in there. What's lacking is a clear description from the person who built the model of what is supposed to be there. He says, "well, this is the overall kill probability and you should think about target acquisition and target selection and weather and terrain" and so on. And somebody says "well, where do I get a number like that?" And pretty soon he's built a little model of his own, or else he's got a little group of people and they find that implicit effect on their own. But somebody has to sit down at any rate and worry about what that implicit effect is. So, we have very real pressures in both directions. As soon as you build that complexity into something, even if it's in a preprocessor, you build a hierarchical structure. And in my experience hierarchies don't seem to be that much more transparent. Although the summary model at the top may be very transparent, the first question that somebody who doesn't like the result asks is, "well, I need to understand better what's going on at the next level down." And pretty soon he has built himself the whole hierarchy in one nutshell again and he has lost the transparency that you thought you had in the beginning. You have at least structured his thought about it into a lot of different nutshells, but he keeps trying to put them back together and see if he can make any sense. In my experience he doesn't really do that very well. We haven't done very well at separating these parts clearly enough in such a way that we can say the structure is really workable.

Let me leave it at that and put on my session chairman's hat. I understand that Sy Dietchman, who was to present one of the papers on the agenda, will not be here. He has been called away on a family emergency. So, our next paper is from Bruce Anderson at IDA, who was scheduled to talk about attrition of air systems modeling, but he told me last night that he wasn't sure he was going to hold exactly to that topic. I said that was all right. I wasn't holding exactly to my topic either.

16 — *Attrition of Air Systems Modeling Overview*

DR. BRUCE ANDERSON
Institute for Defense Analyses

Air systems modeling and the proper staff use of theater models. A description of IDATAM, a tactical air model.

Dr. Anderson: As Bob said, I'm not going to hold exactly to the topic. I'm going to be somewhat more general, as he was somewhat more general. I think it's germane, though, in the sense that what I'll be talking about will carry on the discussions we had yesterday, then concentrate on an air combat application of those discussions, and then have a natural lead-in to this afternoon's discussion of optimization and game theory structures. The reason for this is that there are some applications of the general concepts to air combat that are just cleaner. Not clean, but cleaner — relatively more clear cut in air combat than on the ground where everything is taking place sort of simultaneously right around the same area. The lead into game theory structures follows because aircraft can be used different ways at different times in the war and this leads to questions of how best or how reasonably should one use them. Game theory can shed some light on that.

We heard a lot yesterday about some limitations in the use of theater level models. Larry Low sort of said "well of course, we're not going to use them to predict anything," and then we heard a lot about how they're no good because they don't predict — that theme sort of ran through the presentations. Then some people said "well, really, they can be useful because they sort of predict." I would agree that they can't predict and that you can't just go out and get the right data base. We work hard to get the right data base, and we work on a model to get it right and to get the right answer. Well, that's a general statement maybe everyone would agree with, as everyone agreed that models are an aid to decision makers — they don't make the decision.

Well, what does that mean? You hear people argue that we have to get better data bases, they have to be verified, that they certainly have to be fully supportable if not even incontrovertible. You have to have better models; models have to be credible. Not credible in the sense that Glenn Kent was talking about yesterday — that you put in more units and you get out better results. I firmly agree with that. The models should be, in that sense, credible, but I don't think that is what most people mean when they say credible. They mean things like Lanny Walker had up on that slide when he said, "Does it play intelligence as rank 10?" "How is its organizational structure as rank 1?" In the parlance that is what I think credible means, and I don't think models have to be credible, and I don't think they have to have "good" data bases. By that I mean they don't have to always be credible in every application, and they don't have to have the right data base that someone can't argue with.

Theater-level analyses (Slide 16-1) tend to use judgments more in the inputs and the assumptions in the models to produce analytical results — results that say if this were the case then this would follow. Lower

Slide 16-1 — USE OF JUDGMENT

- Analyses using theater-level models can integrate judgmental inputs and judgmental assumptions to produce analytical results.
- Lower level analyses tend to rely more on analytical inputs and "simplified" assumptions to produce results that need to be integrated by judgment.

level analyses perhaps have better supported inputs, and more engineering inputs, more analytical inputs, and simplified assumptions in the sense of yes/no assumptions. I don't mean these assumptions are easy to make necessarily, an example might be a given flight path; that is a typical path. Well, clearly that is not the only flight path the Russians might use or we might use. But you can look at it and say "that is a plausible flight path" and just ignore it from there on — that simplifies the problem quite a bit. But I think it raises other

problems in how you integrate the results based on one flight path or one set of flight paths, but that being aside you can use those kinds of assumptions and you produce results that may individually be more rigorous, by some definition of rigor but they need to be integrated somehow, and that's probably where you present this piecemeal analysis, and this piecemeal analysis, and this piecemeal analysis to the decision maker. It's my judgment, it's Reiner's mental black box that he looks at it and says "yes, I process these things and this is what I think comes out." I agree with that description, and that's sort of the difference in the analyses. In one, the judgment factor is more exercised in the input; in the other, the judgment factor enters more in analyzing and combining distinct outputs.

We talked about getting the "right answer" in using theater level models. We want to try to do that — build a better and better model and a better and better data base. But I think the theater-level model would sort of fit into one of the four uses I have list-' (Slide 16-2). I labeled the next three all twos, not because they're the same but because I don't want to talk about them. We all have some feeling I think for what those words might mean (2a, 2b, 2c).

Slide 16-2 SOME USES OF THEATER-LEVEL MODELS

(1) "Right answer"

(2a) Net assessment

(2b) Sensitivity

(2c) Insights

(3) Consistency

What I'd like to talk about is using models to test consistency. But the question is consistency with what, and for that I'll need a definition (Slide 16-3). Let me use the word "staff," and let me define it, for the purposes of this talk only, to mean an analysis that integrates results from lower level analyses using judgment, and that it is what I had sort of referred to earlier, the looking at the details and saying, "well, I'm going to put these together using the judgment." Or maybe it is just pure judgment, like a delphi paper or an analysis that presents results from lower level analyses without integration. Here is this result, here is that result, and that is my product. I think there is a reason for selecting the word staff but I don't intend that it mean more than that for this point.

Slide 16-3 DEFINITIONS

Let "staff" mean either

(1) An analysis that integrates results from "lower-level" analyses using judgment (or is pure judgment)

or

(2) An analysis that presents results for "lower-level" models without integration

Let "model" mean a model plus a set of data

Let a model mean a model plus a set of data; as Bob Farrell said, the data can really drive things. It's easy to say, and I'm not going to say anything else other than that "IDAGAM says 'blah'," but it's the data base, and the correct statement may be a data base as processed by "IDAGAM says." That's too hard; it's easier to say IDAGAM, or VECTOR 1, or whatever. But, when I say model, I mean a model with an associated set of data.

Now, getting back to what I mean by consistency (Slide 16-4). I think consistency can test, say a staff versus a model. By that I mean an analysis is made, and it is put together using judgment. A decision maker may look at that and say, "well, while all these pieces seem sort of reasonable, are they consistent? They each may have different assumptions, use different data; I'm integrating them through judgment, or my staff has integrated them, and is that reasonable? Is there any way that I can test that integration to see whether or not more work is needed?" If the decision maker has a model available to him, say even a small sub-set of his staff, he can run the model, or, pay someone else to run it to address similar questions in the model on which his staff has already made recommendations. Where they differ, I choose the results of the staff, as

produced. That is, if you are forced to produce results right away — you know, tomorrow the paper is demanded and that night you have the results — an aggregated theater level model says buy A and all the results of the staff says buy B, then clearly it should be buy B, I think.

Slide 16-4 CONSISTENCY

Consistency can test

A. Staff vs model
  (1) Where they differ,
    (a) Always choose staff, but
    (b) Can show where relatively more work is needed
  (2) Where they agree, assumptions and data of model are "more explicit" and "more quantified"

B. Staff vs staff
  (1) Weapons/studies
  (2) Services
  (3) Countries

But, I don't think that is a reasonable picture usually. It's not reasonable in the sense that you don't always have to make a decision tomorrow, and, when you do, the problems will come up again. When I served in the Pentagon I heard people say "ah, I'm glad that one is over with" several times. But it's not over with. Don't you think you're not going to hear that question again. Of course, you will. Maybe a slightly different word, a different airplane, a different truck, whatever, but the same kind of question will come up again, and that is where these models, are very useful. If you have the time available to say "hey, that question is too hard — we couldn't answer it in a week like we thought we could, or a month, because we've looked at our integration and we want another month to talk about it," or, you can say "this is the best answer we can give now, but in next year's work program, or next year's research budget we can tackle it again." You know, or have a feel at least, where you can relatively allocate budgets. For example, suppose the analyses agreed in eight out of ten cases — your staff analysis and the model. In the ninth and tenth they disagreed. Well, you may be able to say, more on the basis of argument than just gut feeling, that next year I'll concentrate relatively more of my effort in those ninth and tenth areas. That is where I think maybe that model credibility comes in. If in that ninth area you look at the model and can relatively quickly dismiss it and say "my staff came up with a better answer. I think it's highly likely when I look at the details that I'm going to find out that my staff considered more things properly than the model did, and under those circumstances that particular model is not credible." On the other hand, if you look at it and say, "hey, for this kind of data and this kind of structure, the model should have shown a difference. My staff says this thing is twice as good and the model says something else is twice as good. The model is not a perfect model, but it is credible enough and it should at least show something in the same direction." It is credible enough to say next year, or next month, or next week, if you have the time available, that is where you should put in the time as opposed to looking at re-refining to another decimal place the other eight inputs where they agreed.

Of course, where staff and the model do agree there are advantages in that the assumptions and the data of the model tend to be more explicit, more quantified. You can use them for references, for some of the other uses on the list, like sensitivity analyses, and the next study then can refer back to these quantified inputs. But it is very hard to refer back to a judgmental study in which it was a case of "We looked at these problems and we concluded this was best." It's very hard to do anything with that kind of integration in the next study. A model with an associated data base, a documented model, that is the important thing. If we look at a model and it's not documented and it's not available for you to look at — that's not very useful either. That's a judgmental analysis for all practical purposes.

There are other considerations that arise when you compare results from two staffs. The staffs could be weapon systems proponents, or different studies. They could be two services, or could be two countries. Countries we've used a couple of times, but I think the primary use would be comparing recommendations from services, or comparing recommendations from weapon systems groups, or different studies within a service.

It's all very nice sounding, but what are the problems with it? Well, it was suggested to me that a prime problem is obtaining data, obtaining staff data, planning data. That is what I meant when I said data does not have to be good enough. Suppose you're evaluating A-10s. I don't think you care about what the "right num-

bers" are for the attrition of A-10s for the purposes that I'm talking about here. What you do care about is why are we buying those A-10s instead of some other resource, say, A-7s or F-16s, and if we're buying them for some reason, the people who have decided on them have to have some feeling I believe — it's in that black box. But that is the decision maker. He has a staff and the staff has a staff. I'm not so cynical to believe that people don't have at some level a feeling that, well, the attrition rate will be about such-and-such and that translates to a P_K at a certain order of magnitude, and that order of magnitude versus the P_K at the other order of magnitude says that the A-10 is the more cost effective weapon. It's reaching that sense of things and getting hold of it. It's there, and that is the data that counts for the use I'm describing. Not if we had another Arab/Israeli war and they flew A-10s how many would be killed because the Arab pilots are trained better than the Israeli pilots, or vice-a-versa. So, do we even have a consistent rationale for what we're buying (Slide 16-5)?

Slide 16-5 PROBLEMS WITH ATTEMPTING TO TEST CONSISTENCY

A. Obtaining "staff" data
   (1) "Decision-maker"
   (2) Staff
   (3) Studies that use documented models

B. What incentives are there to test consistency?

Well, you can get that by questioning the decision maker, and, of course, that can be very difficult. The Secretary of Defense is a decision maker, but he doesn't have time to sit down and fool around with your model. He is a very busy man. He has an office and his office has a staff and he has a staff and his staff's staff has a staff, and his staff's staff's staff has a staff, I'm pretty sure, the last time I looked at the phone book. But you can get down to a level where there are people who could, if the incentive is there, talk with you. I've sat in the Pentagon for months and talked with the people on generating data base. The incentive was there for those circumstances, and the numbers are there at least implicitly in the judgmental model, and it's not an attempt to trip up somebody. If you can get a consistent set by talking to the decision maker's staff's staff's staff and get some numbers from them, and the next person, and put them together, and can say "hey, here are some consistent decisions and here are some inconsistent ones," I think that is the use, even if you did not get to talk to the personal decision maker directly.

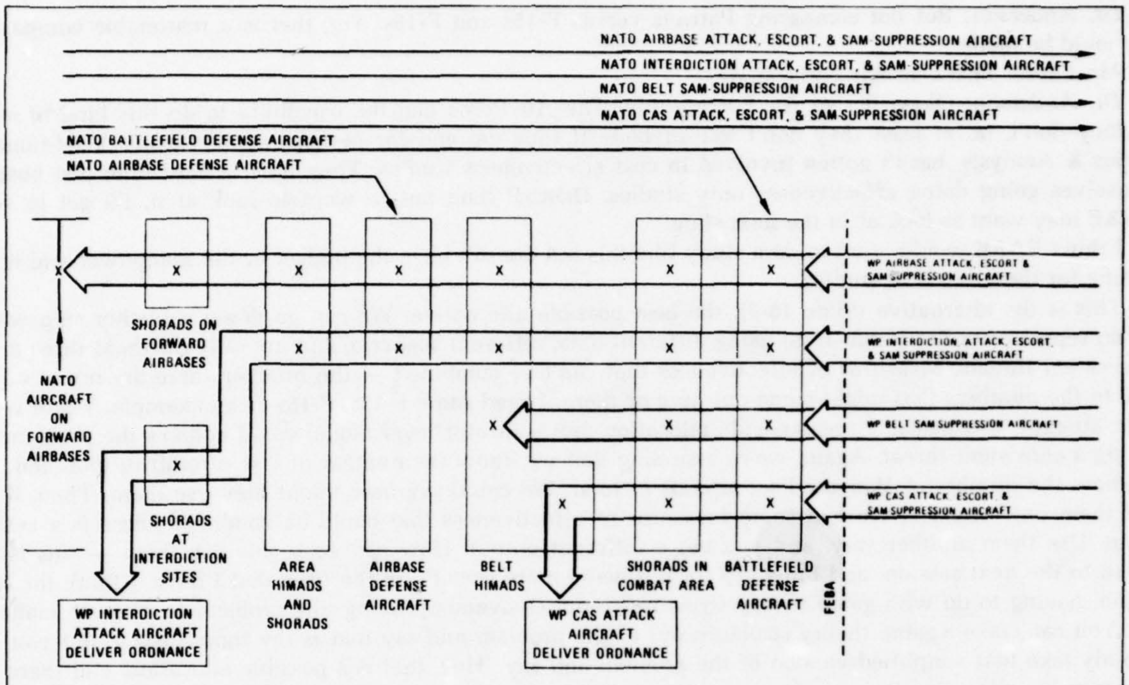
So, you can talk to the staff, as I just mentioned. And, you can get a feeling for the consistency from studies that have used documented models. Those provide a tremendously valuable source. You publish the study, and you build on this information, unlike a judgmental analysis where you really have to start from scratch because you can't reproduce the final integrating judgment that occurs at the top. You can look at previous work if they documented the model and the data base. You can build on that and say, "well, here, this is fine, but this other study justified a similar weapon system and they used that, now they both can't be right. You can't have a P_K of 0.9 and a P_K of 0.1 simultaneously." That is, I think, the real problem — the discrepancies.

If you gave me some money right now and said "test consistency!" This would be a problem in the short range, a real problem. There is no real incentive for a decision maker to test consistency. By the time he's tested consistency he's got a feeling already. He knows what he wants to buy. Why should he want some model to show him that he might be wrong? The incentive perhaps has to come from some person at yet higher level saying "hey, you came to me with the weapon system and it looked pretty good, and then you presented this weapon system and it looked pretty good, and then this one. Individually they all seem reasonable, but how do they all fit together?" This happens occasionally and sometimes the people who are pressing these things will back up and say, "well, we looked at it and it's judgment in the works. Of course, there is an incentive, as sort of an internal rationale, in saying that "I'm really trying to have the best defense for my dollar and I'm trying to help my country," but it's tough to keep in mind that it's got to fit into the planning cycle. The test for consistency doesn't fit in; it doesn't help you make your decisions easier, but it helps to make your decision better. I think there's a very big difference there and that's why there aren't really strong incentives unless they're imposed from someone higher up who asks "are these piecemeal results you're presenting consistent?"

One thing that's been done in lieu of that, and I've talked about staff versus staff and staff versus model, is to look at model versus model. That's not an end in itself, I think the other two are, but I think it serves as a comparison, a surrogate for the other two. I could talk some more about that, but let me go on.

Let me try to be a little bit more specific. This is a general picture of a model (Slide 16-6) that I recently worked on. We call it IDATAM as opposed to IDAGAM, so you have to pronounce it very carefully. Roughly speaking, it takes the air combat model of IDAGAM, let's call that Air Combat 1, and enriches it significantly, let's call that Air Combat 2. The key here is that aircraft can be used in a wide variety of missions. In this case . . . this is half of what really should be the entire chart. If I had John Bode's presentation, I'd have an exactly similar chart right next to it. If you'll notice from all the words that NATO attacked from left to right and Warsaw Pact attacks from right to left, the potential is exactly symmetric. I'm not saying the physical characteristics of the aircraft are the same, whether both sides *would* use the aircraft the same way or *should* use them the same way, but that they *could* use them the same way. So, a full chart would have one right next to it exactly the same. These show the potential allocations.

SLIDE 16-6 IDATAM



X denotes the interactions that can occur between NATO and WP resources on the NATO side of the FEBA (a completely symmetric set of interactions is modeled on the WP side of the FEBA).

Note: This figure is stylized, so it does not give the order in which interactions occur in IDATAM, nor does it include deep attacks or helicopters.

Now, if you look at air combat only, and this is an argument Air Force Studies and Analysis made for awhile, the payoff comes in being able to deliver support in terms of gound forces. In doing that, you may want to fly aircraft to attack enemy aircraft at their bases, or in this case the other side might want to attack yours and you'd be flying to attack his bases. You may want to do surface-to-air missile (SAM) suppression, as shown in the slide or fly escort missions to help you. But the payoff is the attackers, the bomb droppers. You might compare these bomb droppers, then, versus the other bomb droppers. There is sort of a dual role for SAMs; they can sort of shoot at the people forward trying to deliver bombs in the combat area, or at the people going through to the rear.

There is a question here that one might ask. I think it's a reasonable question; it's something we haven't talked about at all in Session I or Session II, and Bob (Farrell) didn't get into it at all. The question is "Are models being used in a cost effectiveness sense?" Jerry Bracken mentioned it briefly in his talk. I don't know how many people in this room have done this sort of thing. We have talked about theater level models, but have we asked this sort of elementary question? This isn't asking how much is enough, but it is assuming we have a fixed amount of money and want to spend it on a certain set of resources. That's a real question and it's really being addressed. But it's not being addressed using theater level models. I think that theater level models could address this. They're not going to provide the right answers, but we're buying certain systems. If there is a rationale for doing so we have to have some feeling for the interactions, the P_Ks etc., the fre-

quencies of occurrence. It isn't entirely made up of future things we know nothing about. The planning data are there; the actual combat data are not. Well, we could use that planning data.

Well, it's not done and nobody's going to do it either. I shouldn't go on that quickly. The Army is going to do it. Phil Louer stated the Army's problem precisely. The Army's interested in what the Army wants to buy.

Mr. Louer: The Army does do something like that.

Dr. Anderson: The Army doesn't even have the capability to do that. I realize quoting things out of context is difficult because, for example, Bob Farrell said IDAGAM is — what was the phrase . . . "The one outstanding model was IDAGAM." Well, one could interpret that several ways. For example, the Army, in comparing its Patriots and F-15s and F-16s, has been using the COMO model, which is a theater-level model by some people's description — if a rose is an F-15, COMO is a theater-level model.

Mr. Louer: I thought you were talking about that type of study. The Army does do that type of study on Army planes.

Dr. Anderson: But not comparing Patriots versus F-15s and F-16s. Yet, that is a reasonable comparison that could be made.

Mr. Louer: That's an Air Force study.

Dr. Anderson: Okay, it's an Air Force study. The Air Force had the capability to do this kind of study but they don't, or, at least they don't tell anybody if they do, and for perhaps good reasons. Traditionally, Studies & Analysis, hasn't gotten involved in cost effectiveness studies. They have a hard time just keeping themselves going doing effectiveness only studies. DDR&E tend not to want to look at it. I'll get to what DDR&E may want to look at in the next slide.

I think PA&E would want to do a study like this but doesn't have the budget or the manpower and is not pushing for these kinds of studies.

This is the alternative (Slide 16-7), the best possible alternative. We can go down a further step where you do separate studies (Slide 16-8) using different data, different scenario, and not only different threats, but inconsistent threats. Measures of effectiveness that can't be combined — the numbers here are not at all related to the numbers that might come out here or there. I used some F-15s, F-16s as an example. There won't be an attempt to consider a theater-wide allocation, but a theater level model could address the problems of making a consistent threat. Again, we're assuming that we know the number of U.S. aircraft in total and that we know the number of Warsaw Pact aircraft in total. We could say how might they use them. Then, if we used them one way and we had some measures of effectiveness that could be combined there is a certain output. Use them another way, and you get a different output. How can each side use them — this is the lead-in to the next session, and I'm sorry for not being more specific in the time that I have. I think the next session, having to do with game theory structures requires oversimplifying the problem to make it manageable. You can't take a game theory structure out of the problem and say that *is* the right answer, but you can certainly take that simplified version of the problem and say "Hey, that is a possible allocation, and there are other possible allocations"

Slide 16-7 FORCE PLANNING ANALYSIS

AWACS	A-10s	F-15s	F-16s	PATRIOTS	SHORADs

ASSUME

(1) WP resources known and fixed

(2) Non-U.S. NATO resources known and fixed

(3) U.S. budget known and fixed

(4) U.S. costs known

I strongly disagree with those who say we should worry more about what we think they will do, read their literature, and see what they plan on doing. However, I strongly agree with Glenn Kent's arguments; other people have made the same argument before, but he's written it down in TAC CONTENDER. He says we ought to try to assess what they *can* do, not what they *will* do. And I think that's right. I haven't heard TAC CONTENDER mentioned as often in the last five years as I have in this room. Five years is an exaggeration, but TAC CONTENDER died away in terms of a way of looking at problems. Slide 16-8 shows the problems on lower levels where the money is being spent in terms of studies and analysis, and I think theater level models can make a contribution in terms of testing consistency by not spending so much money on that kind of analysis and spending more on theater level aggregated analysis.

Dr. Bracken: Bruce, could you make a point about special purpose systems and general purpose systems analyzed in a separative environment.

Dr. Anderson: I think you can put up any set of systems you want of a well scoped study. I could throw in tanks, I could take out AWACs and A-10s and just talk about comparing these. But you have to consider, if you're going to buy an F-15 or an F-16, that it can do various things at various times in the war. Say you want to compare a special purpose system, say a Patriot, versus a general purpose system, say an F-15 or an F-16, and you do this study alone (Analysis 2, Slide 16-8) from a cost effectiveness point of view, or even from only an effectiveness viewpoint, and you then find out how much better a Patriot is than an F-15. But you can't do a cost effectiveness study at this level and come out with any meaningful answer because a Patriot can only do one thing and you're only testing that one thing the Patriot can do, while the F-15 can do two, three, four things, ten things, if you look at that previous list. There is no way you can compare these results and come up with a valid cost-effectiveness conclusion. Reiner?

Dr. Huber: Is IDATAM tied into the land war model or have you some other measure for air effectiveness?

Dr. Anderson: We simplified it again and put it in as part of the air combat model in TAC WAR. But in its full sense, no. All it does is produce as output the number of aircraft by type that fly each mission by type on each day. So it's sort of like the same kind of measure that TAC CONTENDER would produce but in much more detail.

Dr. Huber: Well, does the model also account for, let's say, in the game theory sense for intelligence . . . .?

Dr. Anderson: No, no. What we have done in using it is to input various strategies like "what if the Russians do this . . . .?" We form a payoff matrix by varying some red strategies and varying some blue strategies. Once you do that you've got enough. The difference between that and going to some intelligence agency and saying what will they do — no matter how stupid it may seem in the light of what we do — is so horrendous.

Dr. Huber: No, what I mean is if you're talking about a trade-off between a Patriot and an F-16, for example, you need to do more than count aircraft sorties. You need to reflect somehow the indirect capability of the Patriot to support, let's say, your own aircraft sorties, or to reduce the enemy's capabilities.

Dr. Anderson: Sure, and it would, because Patriot is here in what is called the HIMAD, and it would shoot at enemy aircraft trying to bomb your airbases. And, in fact, in running these studies that capability was very important. I don't want to get into details, but let me say this: you do not give up all of your defensive capability in order to have multipurpose systems because if you have no defensive capability they come in and they destroy your offensive systems. You may be great the first flight, but then you're nothing thereafter. Defense is very important. How do you get that defense and what should the mix be is a different question, but you do count the effectiveness of the defensive system in terms of its ability to destroy these guys up here that are going to hit your potentially offensive systems. Is that the kind of thing you're thinking of?

Dr. Huber: Yes.

Dr. Anderson: Okay.

Mr. Farrell: I have a feeling that several of us might have an active discussion during the panel session with Bruce, but not necessarily concerning air systems attrition.

Our next speaker is Mr. John Underwood, from CNA, who will talk about the attrition of marine systems.

Slide 16-8 POSSIBLE "ONE-STEP LOWER-LEVEL" ANALYSIS

Analysis 1 C^3 Analysis	Analysis 2 Deep Defensive Analysis	Analysis 3 Shallow Defensive Analysis	Analysis 4 Shallow Offensive Analysis
Measure effectiveness and survivability of AWACS	Compare some F-15s F-16s, and PATRIOTS	Evaluate some SHORADs holding PATRIOTs constant	Compare some A-10s, F-15s, and F-16s
Considering data set D_1,	Considering data set D_2,	Considering data set D_3,	Considering data set D_4,
Scenarios S_1,	Scenarios S_2,	Scenarios S_3,	Scenarios S_4,
Threat T_1, and	Threat T_2, and	Threat T_3, and	Threat T_4, and
MOEs $M_{11}, \ldots, M_{1_n}$	MOEs $M_{21}, \ldots, M_{2_n}$	MOEs $M_{31}, \ldots, M_{3_n}$	MOEs $M_{41}, \ldots, M_{4_n}$

17 — *Attrition of Marine Systems Modeling Overview*

MR. JOHN UNDERWOOD
Center for Naval Analyses

Mr. Underwood: Good morning. I guess generalization is the thing for today. Like the others, I'm probably going to be a bit more general than the title of my talk suggests. I think that probably stems from the fact that in twenty or thirty minutes you can't get very deeply into attrition modeling or you find yourself immersed in going exactly nowhere.

Hierarchical analyses and models in naval force planning.

In the next twenty minutes I'll try to describe theater-level analysis as it's being practiced at Navy headquarters. You may notice that I'm referring to this as analysis rather than gaming. I do that because theater-level gaming in the Navy for the most part has taken the form of paper studies in which alternative opposing strategies are examined rather than war gaming in the stricter sense. I sometimes call such analyses analytical war gaming.

I'm referring to a series of major studies conducted by, or sponsored by, the staff of the Chief of Naval Operations. These studies are designed to assist in decisions about Navy force levels and force structure. We've been prominent in the continuing debate about why we need a Navy and what kind of a Navy we need. I don't want to imply that what I'll talk about is the only theater level gaming that's being done in the Navy. For example, the Naval War College has an excellent facility for conducting computer-assisted manual war games. It can accommodate war games ranging from elemental tactical games to theater-level games. It's best and most frequent use has been gaming major naval operations of less than theater-level proportions with a game time duration of a few days. It's been used quite successfully by Fleet commanders, such as CINCLANT Fleet to war gaming these kinds of operations, serving principally to train and educate the staff, to test contingency plans, and to gain insights into strategies, tactics, and capabilities. Some similar gaming and some theater-level gaming analysis has been conducted at Atlantic and Pacific Fleet headquarters. The Naval War College has recently established a Center for Advanced Studies, where some theater gaming and campaign analysis will be done.

Today, however, I'll talk in general about the analytical gaming or campaign analysis that is characteristic of the Navy's major force level and force structure studies of the recent past.

This is a list of the principle force level studies carried out or sponsored by the OpNav staff in the last 15 years (Slide 17-1). The studies each required from 6 months to 2 years to complete and consumed from 4 to

SLIDE 17-1 MAJOR NAVY FORCE LEVEL STUDIES

WAR AT SEA	1960's
CYCLOPS I, II, AND III	1960's
MAJOR FLEET ESCORT	1960's
WAR AT SEA II	1960's
WAR AT SEA (NOW)	1960's
ASW FORCE LEVEL STUDY	1960's
NAVAL REQUIREMENTS AND CAPABILITIES --GENERAL PURPOSE FORCE (NARAC-G)	1971
SEAMIX I	1973
SEAMIX II	1975
SEAMIX III	1975
SEA WAR 85	IN PROGRESS
TACTICAL NUCLEAR WAR AT SEA	IN PROGRESS

25 man-years of professional work. I'll generalize somewhat on these studies, and then describe the methodology that we're using in a current study called Sea War 85. This methodology has evolved through the series of studies as the state of the art has progressed.

Most of the studies on the list you saw were attempts to assess the capabilities of our naval forces in a major war with the Soviet Union in a given scenario or set of scenarios (Slide 17-2). The studies considered scenario variations and examined the effect of variations on several key assumptions such as the length of strategic warning, alliances, and the granting of base rights. In most cases, alternative strategies and tactics for both sides were examined. In some cases, force levels and force structure of the Navy were varied.

SLIDE 17-2 CHARACTERISTICS OF NAVY FORCE LEVEL STUDIES

- OBJECTIVE
  - ASSESS CAPABILITY OF U.S.NAVY
- SCENARIO
  - WAR WITH USSR, VARIOUS SCENARIOS
  - WITH AND WITHOUT ALLIES
  - MOSTLY NON-NUCLEAR--SOME NUCLEAR VARIATIONS
- VARIATIONS
  - MAJOR ASSUMPTIONS
  - STRATEGIES
  - FORCE LEVELS AND STRUCTURE (SOMETIMES)
- OUTPUT
  - ATTRITION TO SEALIFT AND FORCES
  - CAPABILITY TO ACHIEVE THEATER OBJECTIVE
  - RELATIVE STRENGTHS AND WEAKNESSES
  - POTENTIALLY PRODUCTIVE STRATEGIES
  - ROLES AND CONTRIBUTIONS OF VARIOUS TYPES OF FORCES

Measures of effectiveness include the attrition of sealift ships and cargoes and losses of friendly and enemy forces. From these measures, we draw inferences about the ability to achieve theater objectives with a given allocation of forces to the theater. This may lead to conclusions that for a given overall force level, objectives in one theater must be sacrificed to gain a satisfactory outcome in another.

Such studies indentify potentially productive strategies and tactics for both sides. Another important contribution is the insights they yield about the roles and importance of the various types of naval forces. It's only in a broad campaign analysis that the contributions of the various types of forces can be placed in perspective.

Once a good data base has been studied and refined, variations can be introduced to answer many interesting "What if" questions. This has been one of the better uses of this kind of a study. These kinds of variations are possible long after the original study is done, provided the study is documented thoroughly. For example, in a separate study variations were done on the Seamix I study to explore the effect of a significantly larger overall Navy force level. Another study, Sea Express, examined the payoff of various measures to accelerate the sealift of reinforcements to Europe, using Seamix I as a basis.

I'll now describe in fairly general terms the methodology we're using in the Sea War 85 study, a current study in the series that I showed you earlier (Slide 17-3). The analysis is done on three levels — and I'm sure

SLIDE 17-3 THREE LEVELS OF ANALYSIS

- CAMPAIGN
- MAJOR FORCE ENGAGEMENT (MANY-ON-MANY)
- TACTICAL OR UNIT ENGAGEMENT (ONE-ON-ONE)

it's not unlike the other kinds of analyses that are done in the defense establishment — (1) the campaign level, (2) the task level, or, task force, or convoy engagement level, and the (3) tactical or unit level. I'll start with the campaign level and work down, although the analysis is done in reverse order.

As in all war gaming operations, the first step (Slide 17-4) is to flesh out the scenario by clearly identifying assumptions and establishing the national objectives of both sides, then strategies are developed to achieve those objectives. The tasks for naval forces are derived from these national strategies and objectives, and this leads to identification of possible naval strategies. For a given set of naval strategies, we proceed to develop concepts and plans, and to deploy forces on both sides. The logical sequence that I've just described may sound rather elementary, but I find that this is a critical part of the study. If the study is to have any validity, it's important that the objectives and strategies be clearly defined, and that the allocation and uses of forces follow directly from these objectives and strategies. Moreover, it is usually necessary to do a considerable amount of experimentation and analysis during the course of the study before we're satisfied that the forces are used in a near optimal way to support the strategies that are chosen.

SLIDE 17-4 SCENARIO DEVELOPMENT

- IDENTIFY MAJOR FIXED AND VARIABLE ASSUMPTIONS
- ESTABLISH NATIONAL OBJECTIVES AND STRATEGIES
- DETERMINE NAVAL TASKS
- IDENTIFY NAVAL STRATEGIES
- DEVELOP CONCEPTS AND PLANS
- ALLOCATE AND DEPLOY FORCES

The Navy's problem in a major war with the Soviet Union will probably be worldwide. It seems inconceivable that we would be shooting at each other in only one theater when we're in contact in all the ocean areas. Although what happens in any one theater is obviously related to action in other theaters, we normally break the study into analysis of individual theater campaigns being careful to keep the separate parts consistent.

In Sea War 85, we're looking at variations in force levels in each theater campaign, and the theaters, of course, are the Atlantic, the Mediterranean, and the combined Pacific and Indian Ocean theaters. When we combine the results in one way we're exploring the effects of different allocations of naval forces among the theaters. Other combinations of the same results can show the effects of variations in the overall force level.

After strategies are selected and forces are deployed in a given theater, we conduct a campaign analysis in time steps (Slide 17-5). Our time steps are variable in length. During times of intense activity we use short time steps and then lengthen them when the war settles into some kind of a routine.

SLIDE 17-5 TIME-STEP ANALYSIS

- SCHEDULE ENGAGEMENTS BETWEEN OPPOSING FORCES
- MODEL THE SEQUENCE OF ENGAGEMENTS
- RECONSIDER--REITERATE
- MOVE TO NEXT TIME-STEP

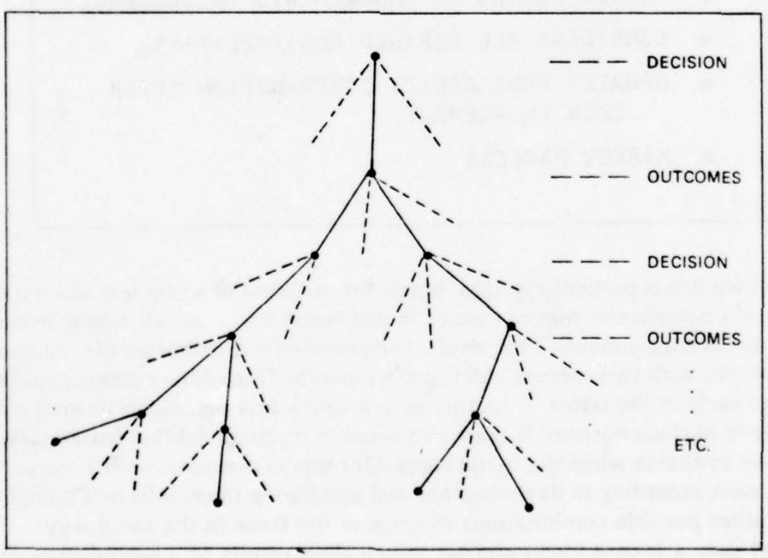
Based on the initial deployments of forces, and the timetable of activity from the scenario, we deduce the engagements that could occur during the first time step and schedule them in sequence. We have campaign models that analyze these sequences of engagements. I'll discuss these models in a little more detail in a moment.

After we model the engagements in a time step we review the results with the advantage of hindsight. We ask ourselves, how could we improve the results, particularly for the losing side. Frequently, alternative tactics are suggested that lead to better results. When we're satisfied with the results of a time step we move on to the next. Decisions in each time step are influenced by all previous results, and, after each time step, we reconsider all of the actions and results in all of the preceeding time steps. In this way we work our way through a base case.

Variations on a base case usually require less detailed attention and frequently they can be automated to a large extent. This process is laborious and time consuming. In theory, one could devise a model that would automate the process, in practice our attempts to do so have resulted in models that are too large and too complex, and the analysts tend to lose control of the analysis. We find that we have a much better understanding of the problem when we walk through it in the manner I described. To automate the process would require that we either simplify the problem to the extent that it would neither be credible nor satisfying, or that we develop a very large complex model with its inherent difficulties. I think that a reasonable compromise is to walk through the base case and some of the variations in the manner I've described and then use a simplified and automated model to find, for example, optimal allocation of some of the major elements of the forces and to crank out some of the variations.

The analytical gaming process that I've described leads to results that are illustrative but certainly not exhaustive. The process may exhaust the analysts but it doesn't exhaust the possibilities. The universe of possibilities could be described by a probability tree (Slide 17-6). As we move through the tree and come to decision points and to branch points, having several possible outcomes, we consider only the paths that appear to lead to interesting results; decisions that we judge to be illogical or clearly disadvantageous are discarded, as are many of the less probable outcomes. Thus, in the final analysis we've only examined a select few of the many possible paths through the tree, but the process does select interesting paths that illustrate outcomes when both sides make logical decisions, and when neither side makes gross errors, and where highly improbable results are ruled out.

SLIDE 17-6 CAMPAIGN ANALYSIS PROCESS



We use several campaign models. One models the campaign against shipping (Slide 17-7). That is, it considers the submarines and air raids that are committed to an anti-shipping campaign. It models the movement of enemy submarines to and from their anti-shipping stations, and their activities while on station. It accounts for their encounters with shipping that sails independently or in convoy. It also models their encounters with ASW

forces in the area, including ASW barriers, area search-and-destroy forces, and the integral defenses of convoys. It treats air raids in a similar fashion. This is an expected value model using a Markov process.

SLIDE 17-7 ANTI-SHIPPING CAMPAIGN MODEL

- MODELS THE SUBMARINE'S ACTIVITY
  - TRANSIT TO AND FROM PATROL AREA
    - --ATTRITION FILTERS
  - ON STATION
    - --ENCOUNTERS WITH SHIPPING
    - --ENCOUNTERS WITH ASW FORCES
    - --WEAPON EXPENDITURE
- KEEP SCORE
- MARKOV PROCESS

Another model handles the encounters of a carrier task group with enemy forces (Slide 17-8). It takes as input the schedule of encounters involving various types and numbers of enemy forces.

SLIDE 17-8 CARRIER TASK GROUP CAMPAIGN MODEL

- MODELS SERIES OF ENGAGEMENTS IN SEQUENCE
- CONSIDERS ALL DEFINED POSSIBILITIES
- UPDATES PROBABILITY DISTRIBUTION AFTER EACH ENGAGEMENT
- MARKOV PROCESS

The Markov model we use is particularly appropriate for this kind of a problem since it calculates in closed form, all combinations of engagements that can occur in this sequence . . . or, all combinations of outcomes that can occur in this sequence of engagements. This results in a probability distribution of outcomes. For example, we might start a series of battles with two carriers and fourteen escorts. One of the carriers could be put out of action by an enemy submarine early in the series. If so, the force is much less capable in its next battle with a force of enemy aircraft using air-to-surface missiles. It makes no sense to try to model that battle using an expected number of 1.6 aircraft carriers available when the battle starts. Our model considers both a one carrier case and a two carrier case, weighting each according to its probability and combining the results in a probability distribution of outcomes. It treats all other possible combinations of ships in the force in the same way.

Still another model uses a Monte Carlo process to calculate results of other engagements in which naval forces other than carriers are involved, and we have another model also that I've mentioned that calculates air strikes on land targets.

These three models operate in parallel to calculate the results during a time step. Both the inputs and the results are correlated and integrated by hand. I would like to emphasize that these campaign models are only bookkeeping models, that is, they just keep score. They do not make tactical or strategic decisions, nor do they simulate tactical interplay. The parameters we use as inputs to the models reflect detailed analysis of both single

unit engagements and task group or task force engagements.

To use these campaign models, it's necessary to have effectiveness parameters representing the probable outcomes of many-on-many engagements and one-on-one engagements. Most of the results representing the many-on-many engagements are derived from previous studies of such engagements. For example, the Navy has just completed a comprehensive study called the Navy Force Mix Study, of both convoy and carrier task group engagements. Since this was a Force Mix study, a wide variety of alternative force compositions in these engagements were examined, yielding results that in many cases are directly usable in our campaign models. A similar study of engagements between other types of naval forces is in progress at the Center for Naval Analyses.

The problem of one-on-one, or unit level engagements, is a continuing one. Estimates of effectiveness of the various types of forces in one-on-one encounters are required for all types of studies and war games. Analyses of such engagements — we call it tactical analysis — is in the nature of basic research (Slide 17-9). A continuing comprehensive program of tactical analysis is needed to stay abreast of the continually changing problem as new weapon systems appear on both sides, new tactics are developed, and our empirical data base becomes more complete.

Until recently the Navy had no such comprehensive program. Each study group was required to gather data where it could find it and to perform the tactical analysis it needed for the study. The result was a lot of duplication, inconsistency, and time wasted on the front end of the learning curve.

About two years ago the Navy established the Sea Control Tactical Effectiveness study, SEATAC for short, at the Center for Naval Analyses. This study is chartered to provide a continuing effort in the principal tactical areas. Its output is a compendium of analyses of tactical effectiveness for the various kinds of forces. This compendium is continually being upgraded. Current work in that study is tailored to provide the necessary inputs to the Sea War 85 study, but we plan to continue the work, revising and updating it so that other studies will have a good base of current estimates and supporting analysis. With continuity, the quality of the work should steadily improve over time, and future major studies should be able to be completed in less time and with much less expenditure and effort.

In summary, I've given you a broad-brush description of theater level analysis or analytical war gaming as it's being practiced at Navy headquarters. This type of analysis contains much uncertainty, and the results fall short of providing all the answers we need. Nevertheless, it's one of the few ways short of actual warfare of getting a handle on force level and force structure issues. These studies have been helpful in providing insights about the capabilities of our forces, about the roles and contributions of various types of forces, and the viability of certain strategies and the nature of our requirements. Moreover, these theater level studies have been important factors influencing decisions about Navy programs made in the Navy staff, in OSD, and in the Congress. We continue to advance our state of the art for conducting such studies. I'm sure that all of us have much to learn from each other as we share our knowledge and experience in seminars such as this. The result can only be to advance the state of the art and to improve our contribution to Government.

Dr. Farrell: Our next paper by Stan Spaulding of Vector Research will discuss techniques for modeling tactical nuclear warfare, and present a different model than TACWAR, with very similar problems I think, as the basis.

SLIDE 17-9 TACTICAL LEVEL ANALYSIS

- BASIC RESEARCH
- AD HOC APPROACH UNSATISFACTORY
  - WASTEFUL AND DUPLICATIVE
  - INCONSISTENT
  - TIME-CONSUMING FOR MAJOR STUDIES
- THE SEATAC STUDY--A CONTINUING EFFORT

18 — Techniques for Modeling Tactical Nuclear Warfare

MR. STANLEY SPAULDING
Vector Research Inc.

Evolution of a theater-level tactical nuclear weapons employment model (Vector I/Nuclear) and its use

Mr. Spaulding: Ed Kerlin's remarks yesterday about the job of deciding how they were going to structure the TAC nuclear model that they worked on struck a responsive chord, because we went through a similar process, and came to some of the same conclusions in general terms of what we need in a tactical nuclear warfare model. I don't need to spend much time on some of these topics because they have been covered before, but we might go over some of those rather quickly and then concentrate on some issues that weren't discussed, which I think are important (Slide 18-1).

Slide 18-1 — OUTLINE

Background

- Nuclear forces (delivery systems and weapons)
- Employment doctrine (including constraints)
- Combat processes affected by tactical nuclear weapons employment

Modeling Conventional/Tactical Nuclear Warfare

- Specification of nuclear weapons employment parameters
- Target acquisition and allocation of weapons to targets
- Damage assessment
- VECTOR-1/NUCLEAR MODEL
  - — Description
  - — Experience
- VECTOR-2/NUCLEAR

Briefly, these are the kinds of things that you have to deal with when you're talking about nuclear weapons and nuclear delivery systems (Slide 18-2). With respect to cannon artillery, although we don't currently credit the Soviets with having such, we think they're certainly capable, if they want, to put in the 152 gun howitzer. When we talk about tactical nuclear warfare, the large yield weapons, due to various constraints at least on the friendly side are not usually considered, but they are in the stockpile. The evolving technology of nuclear weapons gives you a lot of options in some of the newer versions with dial-a-yield, enhanced radiation to get kills of people in tanks instead of destruction of property and so forth. So your methodology has to account for these kinds of variations in the performance parameters of the systems.

Now, a few remarks about evolving nuclear doctrine: there's a large ferment in the analytic community about looking at nuclear doctrine. Until recently it's been essentially the same for maybe 20 years. Now people are coming to the realization that maybe we ought to reassess it. So the Defense Nuclear Agency has sponsored a large program to study a variety of issues, and the Army has done a considerable amount of work in looking at its tactical nuclear doctrine. I don't know whether it's published yet or not, but there's a

Slide 18-2 — NUCLEAR FORCES

Delivery Systems

- Cannon artillery
- Surface to surface missiles
- Free rockets
- TAC/AIR

Weapons

- Wide range of yields (subkiloton to megaton) in stockpile
- Available weapons may include
  - — Enhanced radiation designs
  - — Reduced residual radiation designs
- Many weapons have dial-a-yield option

draft FM-100-5-1, and some of the points in it (Slide 18-3) include a concept of a package, a corps-size package, to be released within a time frame of 12 to 24 hours and the actual delivery to be in a pulse of not more than about an hour and a half. Part of the idea behind this is to promote the perception on the other side that we are intentionally limiting. Now, one of the problems with all this is when you start looking at the two-sided thing the question arises about what the other side is going to do in response. And, at least in the unclassified literature, the Soviets don't think much of the idea of limited response. If we use any tactical nuclear munitions they're going to use a lot, basically, or at least that is what my understanding is of what I've read of their writings that are available or that I can talk about. One other thing that is kind of interesting that has recently been talked about is the idea of trying, within various constraints including collateral damage constraints, to maximize the actual coverage to try to get additional bonus effect. This is somewhat of a change from some of the earlier ideas of minimizing the yield to get a required effect on the target.

Slide 18-3 — NUCLEAR WEAPONS EMPLOYMENT DOCTRINE
(Based on Draft FM 100-5-1)

Tactical Nuclear weapons provide battlefield support for conventional/nuclear operations

Conventional forces slow enemy and force him to mass prior to nuclear weapons employment

If nuclear weapons are employed, plan is to authorize use by individual corps-size "packages":

- To be released within a "time-frame" of 12-24 hours
- Weapons delivered in a "pulse" of 45-90 minutes
- National command authority (NCA) will limit time span to promote perception by enemy that employment was voluntarily limited
- NCA may specify constraints on:
  - — Collateral damage
  - — Total number of weapons
  - — Maximum yield
  - — Total yield
  - — Etc.
- Within NCA contraints corps may refine package for best tactical advantage
- Maximize lethal coverage

I think that's enough about that topic. These are some of the things you need to consider if you decide to say nuclear weapons have been employed (Slide 18-4). They're going to affect what is happening in a conventional game and that is why you're using them. You want to affect the results of the conventional combat, and these are the kind of processes that will be affected. Now, most of our models including VECTOR 1-NUC, don't cover all of these effects. Specifically, we don't do very much about communications or logistics; we concentrate on firepower and to some extent on command control effects of the nuclear weapons.

Slide 18-4 — PROCESS OF CONVENTIONAL OPERATIONS AFFECTED BY NUCLEAR WEAPONS EMPLOYMENT

Firepower	Intelligence and target acquisition
• Ground-to-ground — Direct fire systems — Indirect fire systems	Communications Logistics
Command and control	Movement
• Mission assignments	
• Allocation of resources	
• Etc.	

In the games that we played, the effects on target acquisition and intelligence wouldn't have too much influence because they were essentially pulse-type situations where we'd already acquired the targets before the weapons were employed in a longer duration exchange — one exchange followed by another exchange, followed by another exchange — and eventually we're going to attrite a lot of our acquisitions resources and that will perhaps have a significant effect. Acquisition, by the way, is a very limiting factor on the utility of your nuclear force, or at least that's one of the conclusions we came to.

We considered the movement of the tactical forces in the vicinity of the FEBA. Primarily we did not analyze the effect of movement of logistics support in the rear area, and that would be something you'd want to do in a more extensive model. When you're talking about the employment of nuclear weapons these are some of the factors that need to be accounted for in some way (Slide 18-5): the time of the weapon delivery, delivery system errors, reliability, the energy spectrum, and the latter has to do with such things as enhanced radiation weapons, whether you're going to get most of your yield in terms of the radiation flux or the conventional or the normal type of spectrum of energy. Variations from nominal yield for some weapons can be

Slide 18-5 — SPECIFICATION OF EMPLOYMENT PARAMETERS

Time	Delivery errors
Delivery system and weapon reliability	• Horizontal
Yield and energy spectrum (e.g., ER, etc.)	• Vertical
Variations from nominal yield	Target location errors
Height-of-burst	

a problem, although that's probably not a major effect. Height-of-burst, planned height-of-burst, and the delivery errors, in terms of horizontal and vertical and target location errors can have an effect. These are all factors that will impact on the particular weapon versus target deployment, and, of course, the target characteristics, the size of the target, the density of elements in the target and the response of the individual target elements to the various effects.

I'm not going to say much about nuclear effects because I think a lot of work has been done that people know about. DNA spent millions of dollars on effects work, and I think there are other issues when you talk

about the theater level campaign that we don't know nearly as much about, that are more important in terms of having a model of the impact of nuclear weapon employment on the outcome of a theater campaign. So, that's all I want to say about nuclear effects per se.

I would like to say something about target acquisition and response time (Slide 18-6). I think this is one of the areas that imposes serious limitations on the effectiveness of the TNF. You've got to have a target located, you've got to get the round on the target before he goes someplace else if you're going to do much good with these things. Currently we don't do very well in this area, and the deeper the targets the more difficult time we have for firing, particularly the more difficult time we have in processing the acquisition information and making the decision and getting a round or a bomb or whatever kind of delivery system we're using onto a target. Ours is not the only analysis that indicated that there are serious problems in that area.

Slide 18-6 — TARGET ACQUISITION/RESPONSE TIME

Significant constraint on usefulness of nuclear weapons

Response time (time from acquisition until nuclear weapon can be delivered) may be long relative to stay time of target

Conventional/nuclear modeling should specifically account for

- Probability of acquiring targets as function of
  - Target category
  - Range from FEBA

Capabilities of acquisition systems

- Probability distribution of target stay time
- Probability distribution of response time, including:
  - Time to process acqusition information
  - Decision time

In terms of allocation (Slide 18-7), this is one of the things that the model must handle in some way — how you're going to decide what targets to attack and what weapons you're going to use to attack the various kinds of targets, and there are lots of constraints. The authority to use the weapons is going to come down with a specific set of constraints. Those must be satisfied. You're also going to be constrained by the number of delivery systems, the number of weapons of various categories you have, the deployments of these things, their ranges — all these things will be constraints on the utilization that must be accounted for in some way in your model. That's kind of a very sketchy idea of the specifications for what a model should do.

Slide 18-7 — ALLOCATION OF WEAPONS TO TARGETS

Must satisfy NCA constraints

- Time frame
- Pulse duration
- Collateral damage
- Maximum number of weapons
- Maximum Yield
- Maximum total yield
- Etc.

Other constraints

- Delivery system availability
- Range of delivery systems
- Deployment of delivery systems
- Weapons available for each delivery system
  - Yields available for each weapon
  - Height-of-burst options
- Targets available (acquired or suspected)
- Target deployments
- Target permanence
- Etc.

Now we're going to talk about the VECTOR 1 (Slide 18-8) nuclear model that we lashed together and what its characteristics are and then discuss very briefly some of our experience with running that model.

Slide 18-8 — VECTOR-1/NUCLEAR

- Combined conventional/nuclear warfare campaign model developed by integration of VECTOR-1 and UNICORN models
- VECTOR-1 features:
  - — Theater-level, combined arms, dynamic, two-sided conventional combat model
  - — Seven combat sectors
  - — Represents front line and reserved maneuver units, weapons, personnel
  - — Battalion level resolution, system on system effects
  - — Represents CAS, artillery, AH fire support and other tactical air missions
  - — Attribution of attrition over time
  - — Used in numerous studies
- Unicorn functions:
  - — Allocates nuclear fires to large target arrays considering many weapon-target capabilities, using mathematical programming techniques
  - — Assesses damage (radiation, blast, thermal) to individual targets for any indirect fire weapon and tactical air
- Interface procedure:
  - — Deploys VECTOR-1 surviving forces into standard target arrays (down to company sized targets) for use by UNICORN and determines map coordinates
  - — Incorporates nuclear only delivery means and storage sites into the target arrays
  - — Aggregates company weapon and personnel nuclear-fire survivors into VECTOR-1 battalion task forces
  - — Assess conventional attrition of TNF and bookkeeps all TNF survivors
- VECTOR-1/NUCLEAR features:
  - — Represents low intensity employment of nuclear weapons in combined conventional/nuclear conflict
  - — Reflects immediate attrition and behavior (shock) effects of nuclear fires on maneuver unit movements
  - — Can be used to reflect localized use of nuclear fires in breakthrough areas down to battalion sized units and company sized targets within them
  - — Permits assessment of effects nuclear fires have on subsequent conventional combat results
  - — Model geometry can be overlayed with real map to utilize real population densities for calculating collateral damage

VECTOR 1 is a combined conventional/nuclear warfare campaign model. It was devised by essentially modifying the UNICORN model, which is a target allocation and damage assessment model. UNICORN is basically what you might call a salvo model; it looks at a kind of a one-shot thing, over a period of maybe several hours, of a set of weapons delivered against a set of targets, and you can input to that various parameters that will tell the thing how many weapons it can use, what is the maximum total yield it can use, what is the maximum yield, and so forth. One of the nice things about UNICORN is that it has convenient ways of making sure that you satisfy all those constraints that we talked about. Another feature that we like about UNICORN is that it's kind of compatible with the level of detail of the VECTOR 1 model. Actually, it's a little more detailed, and we did have to do some disaggregation of the elements in the VECTOR 1 model as I'll discuss in the interface methodology, but it is sort of an expected damage or an assured damage type of calculation. It does not do Monte Carlo calculations. You may have heard of the ATLAS/SATAN lash-up, which was a fairly obvious example of a mismatch in terms of modeling approaches. ATLAS is fairly gross and SATAN is fairly detailed. ATLAS is deterministic, SATAN is probabilistic, and for a variety of reasons we didn't want to go that way. We really wanted something that would allow us to investigate a lot of different cases. Part of the reason for that is uncertainty about what the Soviet response is going to be. We'd

like to look at a variety of Soviet responses and to do that you need something that runs reasonably fast. Either that or you need lots of money, more money than we have.

Okay, VECTOR 1 you've heard about. The UNICORN functions were basically to allocate weapons to targets. It has a linear programming scheme which turns out to be, to us anyway, one of the undesirable features of the model. It eats CPU time and it's a little cumbersome to handle to get it to do what you want it to do, but basically what it does is figure out feasible target/weapon delivery system pairs, and these are the decision variables in this mathematical programming scheme. Then you have target values and you can specify various constraints and each target can have three kinds of value, a target value period, a mobility value, or a firepower value. The optimization scheme tries to optimize the target value destroyed and still satisfy the constraints. However, you can say you want so much of the mobility value destroyed or so much of the firepower value destroyed, and there are various other ways that you can manipulate the model to get it to do what you would have done if you'd had the chance to do it by hand. That's basically all we did, you see, and so it was a roundabout way to get an allocation that we'd already figured out ahead of time we wanted, and for that reason, since the algorithm tends to eat CPU time, it was a little bit of a pain in the neck. But it did at least satisfy all of the constraints you put in it and that is a nice feature.

In addition, the model assesses damage. It takes into account system delivery reliability, weapon reliabilities, target location errors, possibility that the target may have departed before the system or the round gets there, and you get, I think, a reasonably good estimate of the expected damage or the assured damage that you're going to get on targets, and you can specify a lot of variables such as the required level of damage to consider a target destroyed and things of this nature which are of interest to the nuclear folks.

Okay, the interface procedure; one of the things the interface procedure had to do was to disaggregate to some extent the Vector organization. VECTOR 1 allows elements only down to battalion level and we wanted smaller elements as targets. Now, basically, what we had was the consideration of the platoon as the target element on the Blue side, since their defense was spread out. On the Red side, partially because the Soviet's company is smaller, and also because they're on the attack, we considered a company as the target. Well, VECTOR keeps track of Blue battalions and Red regiments, so we had to disaggregate the battalions and regiments into platoons and company target arrays on the Blue and Red sides, respectively. That's one of the things the interface methodology did, and then it fed this information to the UNICORN model. The UNICORN model destroyed certain of the targets, but only on one side at a time because, for a Red strike or a Blue Strike, one could be followed by the other, but at one time you're only considering one strike by one side. Then UNICORN would come back with the amount of the various categories of targets destroyed, and the interface procedure would then convert that back into a modified set of statuses for the VECTOR 1 model. In summarizing the features of the combined model, it reflects what might be called the low intensity end of the nuclear spectrum, although if you were involved in it you might not think it was low intensity. It reflects the immediate attrition. We tried to model some of the behavioral aspects. One point I wanted to make is that I think the behavioral aspects of the response of the units that are hit and the units in the vicinity of units that are hit is one of the things that we don't really have a handle on, and I think that the modeling in this area needs to be flexible enough to reflect a variety of assumptions about this kind of response.

A similar area would be the response of the civilian population in the area. Our friends from West Germany probably have a very active interest in that particular area. What's going to happen to all those people if one side or the other starts using the nuclear weapons? Probably they're going to try to evacuate, clog up all the roads, you know, and create a tremendous impasse. We don't really know what's going to happen there, and I think that when we try to model theater-level nuclear war we need to consider a wide variety of possible responses, and at least make the provisions in the model to reflect a variety of assumptions.

Dr. Kerlin: Do you use this civilian population as a constraint to firing the weapons, and if so, how?

Mr. Spaulding: Basically, the UNICORN model allows you to specify a collateral damage constraint. You can have so many areas where you preclude collateral damage and then in the vicinity of various targets you have density of the population in the vicinity of that target. The model then computes a collateral damage number and a number of civilian casualties, the allocation scheme will not allocate in such a way to exceed an input maximum level of collateral damage.

Dr. Huber: It actually limits yields then?

Mr. Spaulding: Well, it limits yields to the number of weapons. There may be certain instances in which you simply can't use a weapon in a particular place because it's too close to an area where you're precluding collateral damage. To be frank, the UNICORN methodology of collateral damage is not too good, and I think this is a feature that we would want to improve in subsequent work.

Let me talk a bit about our experience with VECTOR 1 (Slide 18-9). We did use it in a study for DNA, in

the capability of a tactical nuclear force to support an event. The results of this were very interesting; if we had more time and if the results weren't classified we could talk about them.

Slide 18-9 — EXPERIENCE WITH VECTOR-1/NUCLEAR

Used in VRI study for DNA, "Analysis of Tactical Nuclear Force Capability," (classified study)

Approximately 300 model runs were made for above study

Investigated sensitivity to

- Adequacy of defense criteria
- First use initiative
- Neutralization recovery rate

Most troublesome model feature was unicorn allocation routine

- Uses "target value destroyed" as objective function
- Getting desired allocations involved manipulating various constraints
- Optimization algorithm eats CPU time

But one of the nice things about the study was that the system allowed us to make a lot of runs, something like three hundred runs total, to investigate this wide range of response. We did have some capability to look at behavioral response. We had a factor that was input, that said that if such-and-such a fraction of the units in either the Blue battalions or the Red regiments were destroyed, the entire battalion or regiment would be neutralized (Slide 18-10). Another input was a recovery rate. Neutralized units were assumed not to be available immediately after the strike, but some fraction of them might come back the next day, and a larger fraction the day after that, and a still larger fraction the day after that, and the fourth day they were 100%. Now, we varied those fractions from 0,0,0 up to 1,1,1; 1,1,1 said there was no effect on the engagement at all the next day, and the base case was, I think, one-half, three-fourths, and one, which were the fractions of the neutralized units that were available on subsequent days. There was some impact from the latter fractions although it wasn't quite as dramatic as you might think. Finally I've already mentioned to you how the optimization routine for target value destroyed can be a little troublesome as it proved to be in this study.

Slide 18-10 — NEUTRALIZATION RECOVERY IN VECTOR-1/NUCLEAR

Red companies (Blue platoons) in regiments (Blue battalions) not destroyed may be neutralized if specified fraction of other companies (platoons) in same regiment (battalion) are destroyed.

Model has variable input specifying cumulative fraction of neutralized units which have recovered 1, 2, and 3 days after strike.

Base case:	$F_1 = 0,\ F_2 = 1/2,\ F_3 = 3/4$
Variations:	$F_1 = F_2 = F_3 = 0;$ $F_1 = F_2 = F_3 = 1$

I'd like to make two points here. First, we ran this model (VECTOR-1/NUCLEAR) primarily in a quasi-interactive mode because we wanted to make sure that the tactical decision rules that we were writing for this thing made sense — Seth talked about the tactical decision rule package. So we had a man in the loop to the extent that we finally got a set of tactical decision rules we thought were good, and then we made a lot of runs with that set of tactical decision rules, without having a man in the loop. So, you can have something in between; a man in the loop, and completely automatic.

Second, there is the idea of hierarchy, and that is, for common situations, you may have table look-up, and I think this would be a good idea. Perhaps you might want to have detailed computations for a lot of cases but for the very common cases of, say, attacking a Red tank battalion with a Lance, you can just table look-up and not spend a lot of time computing.

But, all in all I think the most critical problem is behavioral response.

Colonel DuPuy: You say you considered recovery of units after neutralization. Did you consider delayed radiation effects?

Mr. Spaulding: No, we did not. The VECTOR 2 proposal would have incorporated that (Slide 18-11). I think that it might be an important effect. In the games we were playing which essentially called for a pulse on one side followed by a pulse on the other side, I don't really know whether consideration of those people who'd gotten enough radiation to make themselves sick would have been useful. How useful are those people going to be anyhow? Basically our model considered that even though you didn't kill all of the unit you hit 30 or 50%; they were destroyed and they weren't in the game any more. But that is something that is worthy of investigation.

Slide 18-11 — VECTOR-2/NUCLEAR

Extension of VECTOR-2 theater level model

Modeling improvements in VECTOR-2

- Variable level of resolution
- More flexibility in numbers and types of systems modeled
- More realistic battlefield geometry
- Extension VECTOR-1 "tactical decision rule" package to explicitly portray command hierarchy
- Explicitly accounts for imperfect intelligence, communications delays, decision time, etc.

Allocation of weapons to targets will use extension of tactical decision rule module

Damage assessment will be expected value using standard methodology (e.g., EM-1)

Fallout effects will be portrayed

Dr. Farrell: Now, we have Mr. Ellwood Hurford, from the U.S. Army Logistics Center, who will present his paper on Logistics Support and Combat Unit Effectiveness.

19 — Logistic Support and Combat Unit Effectiveness

MR. ELLWOOD HURFORD
U.S. Army Logistics Center

Army logistics modeling efforts at the theater-level including the MAWLOGS, LOGATAK and MASC models.

Mr. Hurford: As Bob mentioned, Wilbur Payne had something to do with the planning of this session, and I'm in the Training and Doctrine Command (TRADOC) as Wilbur Payne is now. As you know, he runs TRASANA, the TRADOC Systems Analysis Activity. I happen to be from the set of lovers down at Fort Lee known as the Logistics Center. We used to have in the Army, you know, a fairly well-defined outfit. We called the combat types, the fighters; and we had the combat service support types, these are generally the supply, maintenance, transportation, personnel types, they were called the lovers; and, then we had that body of real lovers who used to like to masquerade as fighters; they called them the combat support types, the engineers and signal people and so on. Well, I'm from the element of the

Training and Doctrine Command known as the Logistics Center. We're mainly supply, maintenance, transportation types. TRADOC also has the Administration Center, which is located at Fort Benjamin Harrison, and, by their title, as you can imagine, they have the adjutant general function and the chaplain function and they interface with the medics. Basically, I make that point because, as Stan Spaulding said in his last presentation — he didn't know he was giving me a key — but he said there are a couple of items down toward the bottom of his slide indicating that VECTOR 2, or VECTOR 1, didn't pay too much attention to logistics. Dr. Wilbur Payne asked me to talk about what we're working on as a part of the overall Theater Nuclear Force Survivability study (TNFS). TNFS is a large program being run by Dr. Payne for TRADOC. The TRADOC Logistics Center is responsible for the logistics support analysis in that study. I'd like to tell you today about some of the things that we're doing in the theater-level modeling of logistics.

When I leave home with a bunch of slides, my young compatriots ask, "What if somebody asks you detailed questions about what you're doing?" I respond, "Well, you fellows forget, I studied under Bob Benchley." Now, some of you are old enough to remember Bob Benchley." He used to lecture on briefing techniques. I had occasion to want to use his guidance, but I couldn't quite, when, in 1967, I was briefing General Harold K. Johnson, Chief of Staff of the Army, and I displayed a slide and he said, "I don't believe that slide." I wanted to say, "Well, it wasn't important anyway," because that's what old Benchley used to tell us to do. I also have developed other techniques since 1943 when I graduated from the Coast Artillery Officer Candidate School. I wound up as a second lieutenant, briefing officer. An old Coast Artillery colonel got after me right away with the direction, "Son, when you put a slide up there, don't read it to the audience. They're smart enough to read it." I responded, "But, Sir, I don't know how long to leave it up there to make sure everybody has time to read it." He said, "Very simple, son, pick out an old field artilleryman in the audience and when his lips stop moving, you'll know everybody else had a chance to read it." So, now Colonel DuPuy knows why I had to establish the fact that he spent many years as a field artilleryman.

Enough of that, let's get on with the business. Wilbur said tell them about what you're doing in TNFS. We at Fort Lee are worried about arm it, fuel it, and fix it, and our friends at the Administration Center, Fort Benjamin Harrison, are worried about manning the system. I'm going to speak about a theater-level model that we are using to try to approach the logistics problems of arm it, fuel it, fix it, and, to a degree, man it.

The starting point for the Logistics Attack Model (Slide 19-1), LOGATAK, is the system called MAWLOGS — Model the Army Worldwide Logistics System. What is MAWLOGS? Well, MAWLOGS isn't a model. It's a set of models developed originally by the Research Analysis Corporation. Later, MAWLOGS developments have been undertaken by the BDM Corporation. MAWLOGS consists of modules numbering in the hundreds. Our analysts can assemble a model out of these various modules. We try to model maintenance and supply and transportation through the various echelons by tracing items and actions with a high level of detail.

Slide 19-1 — LOGATAK STARTING POINT — THE MAWLOGS SYSTEM

MAWLOGS (Models of the Army Worldwide Logistic System) is an automated system developed by the army which assembles simulation models of logistic systems tailored to a particular study.

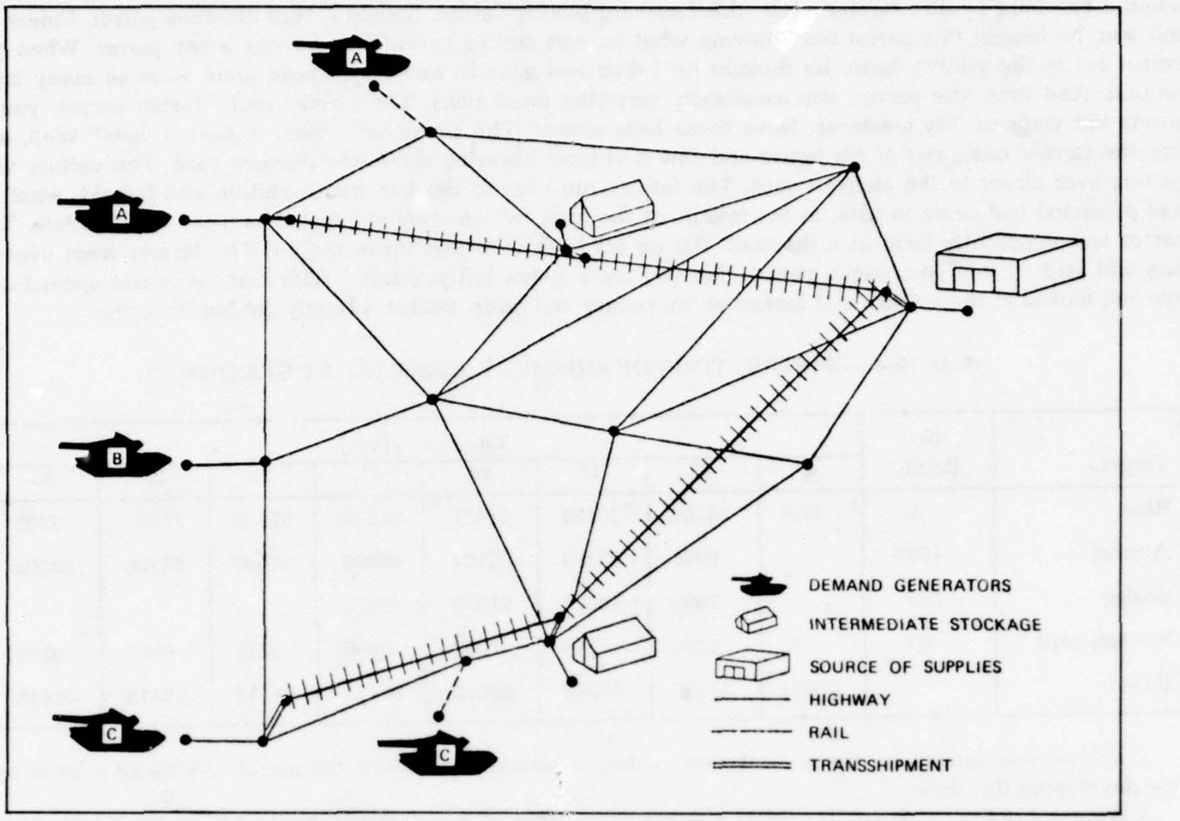
New model — LOGATAK

LOGATAK (Logistic System Attack Simulation) is an extension of the MAWLOGS system under development by BDM/DNA to include the ability to simulate attack destruction and the response of the logistic network to the imposed alterations.

Now, to get to what Wilbur wanted me to talk about, the LOGATAK model. LOGATAK is a logistics simulation. (Slide 19-2). The BDM Corporation developed LOGATAK initially for Defense Nuclear Agency. They took our MAWLOGS model (when I say "ours", it belongs to the Army) and adapted it to the European theater. Although I show a simple schematic, the first and largest effort BDM undertook was to model the European transportation network. BDM modeled the roads, the railroads, and the supply points. The users in this model were the combat divisions.

The model permits the use of up to twenty demand generators. The combat divisions provide the demands on the logistics system. A scenario provides initial requirements. The scenario used in the TNFS

SLIDE 19-2 LOGATAK ACTIVITIES



LOGATAK analysis was developed at Fort Leavenworth using models such as the JIFFY game, the CEM, and DIVWAGS. The specific scenario modeled has the title Europe I, sequence 2A (Slide 19-3).

Slide 19-3 — LOGISTIC SYSTEM ATTACK MODELING REQUIREMENTS

- Evaluation of logistic system interdiction
  - — Location of attacks
  - — Timing of attacks
  - — Effect of interdiction
- Range of scenarios and transportation networks
  - — Movement of demand generators (divisions)
  - — Varying demands for materiel
  - — Effect of intermediate stockage
- Automated optimization aids
  - — Aid to determine best attack strategy

I mentioned the fact that we have to model the transportation network, and the stockage points and the initial stockage. The model produces outputs on the transportation workload and demand satisfaction from the supply standpoints. The new dimension we are now adding in our modeling is the idea of attacking the system. The slide describes the attack modeling requirements. I'm not going to be able to say a great deal in depth of how that works. Instead, I will show an unclassified slide, which illustrates how the LOGATAK

model can be used to analyze logistics vulnerabilities (Slide 19-4). The analyst who presented that LOGATAK example said that it is a big user of data. Most simulations ... That reminds me of the story that I heard about a Leesburg poultry farmer. Now, this Leesburg poultry farmer bought a very amorous parrot. I mean to tell you, he bought this parrot not knowing what he was getting except that he was a boy parrot. When the parrot got to the poultry farm, he thought he'd died and gone to heaven because there were so many hens around. And Pete, the parrot, was constantly servicing these hens. The farmer said, "Listen parrot, you're gonna kill yourself. My goodness, leave those hens alone." The parrot said "Beat it, beat it, bub." Well, one day the farmer came out of his house and saw a vulture hovering above the chicken yard. The vulture was getting ever closer to the chicken yard. The farmer ran over to the hen house and, lo and behold, what he had predicted had come to pass. In the middle of the hens, which were exhausted as usual, lay old Pete. The parrot was apparently dead as a doornail, flat on his back with legs up in the air. The farmer went over to him and said: "I told you, stupid parrot, that you were gonna kill yourself." With that the parrot opened one eye and looked at the vulture and looked at the farmer and said: "Beat it, I finally got her hovering."

Slide 19-4 — METRIC TONS OF SHIPMENT ARRIVING AT STATION

Targets	No. Bombs	Time (days) 9	12	15	18	21	24	27	30
Base	0	3366	23164	23164	41417	51040	67136	78305	78305
Airport	108		8965	23164	23164	45089	51040	67136	78305
Bridge	288		2693	12072	21970	40377			
Storage yard	33	52	2225	2225	7092	8000	8212	9892	9892
(Lost)		22603	38387	53608	65610	68413	68413	68413	68413

Anyhow, simulations require a lot of data — they're insatiable like Pete the parrot. We spend a lot of our time developing this data.

During a briefing at DNA, the BDM representative told of how LOGATAK had been used to analyze deployments. In the illustrated hypothetical case (Slide 19-5), the baseline represents a force equated to metric tons. The attacks may be accepted as the amount of effort involved in attacking the items illustrated. It will be noted that the logistics system can recover from attacks that destroy bridges or airports, but it does not recover from attacks on storage yards during the period under consideration.

Slide 19-5 — TYPES OF ATTACKS — AMMUNITION SUPPLY

Baseline	No logistic attacks
Attack 1	CSA supplies 80% destroyed
Attack 2	CSA supplies 80% destroyed
	Miesau supplies destroyed
	Rhine river line of interdiction
Open bridges	Attack 2 with incomplete line of interdiction

The work done by BDM for DNA appeared to have direct applicability to the efforts underway at the TRADOC Logistics Center in the TNFS project. Accordingly, the BDM Corporation accomplished what has been called a reconnaissance using the LOGATAK model to assess the vulnerability of the ammunition and POL distribution systems in Europe. This slide lists the types of attacks on the ammo distribution system. Note that Attack 2 is the most severe of those listed.

Slide 19-6 illustrates the status of ammunition at the corps storage areas. Note that under Attack 2, the percentages of demands filled drops to unacceptable levels. The combat units would not be resupplied with required ammunition.

Slide 19-6 — CORPS STORAGE AREA (CSA) AMMUNITION STATUS (DAY 4.0)

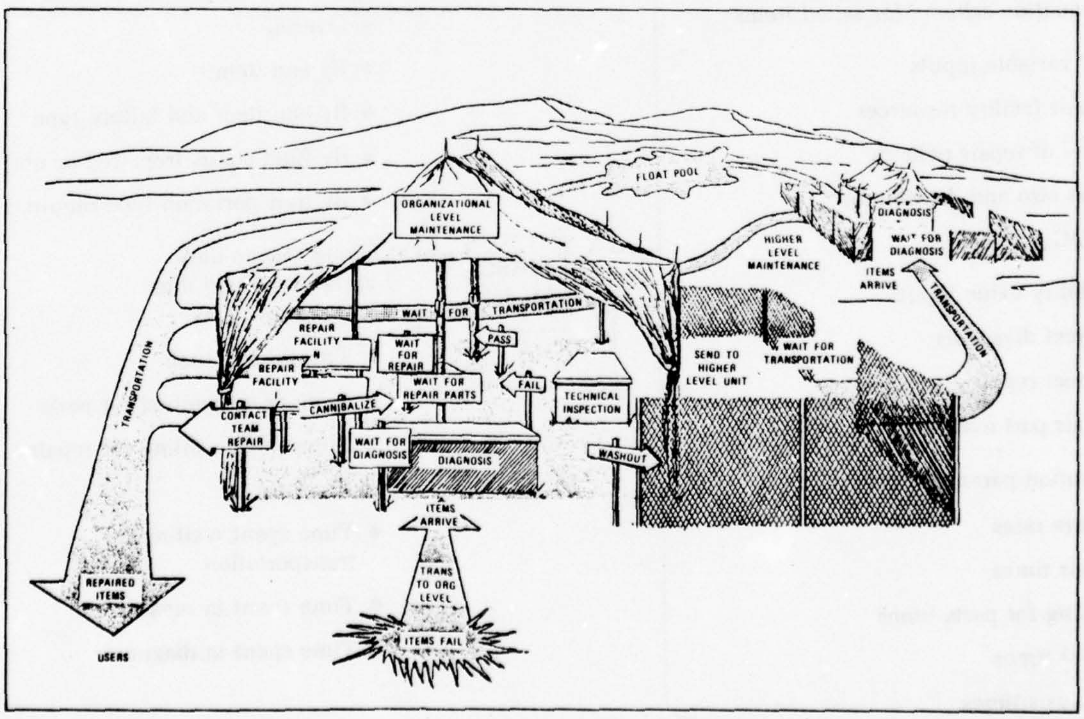
Attacks	Demands Received		% Demand Filled		Average Balance on Hand		Final Balance on Hand	
	CSA01	CSA02	CSA01	CSA02	CSA01	CSA02	CSA01	CSA02
Baseline	11,447	10,373	100%	100%	87%	79%	96%	92%
Attack 1	11,447	10,373	93%	86%	71%	63%	96%	92%
Open bridge (Day 2.0)	11,447	10,373	89%	78%	39%	28%	90%	71%
Attack 2	11,447	10,373	35%	18%	1%	1%	7%	0%

The LOGATAK model is being used to evaluate supply and transportation vulnerabilities. At present, we do not have a model which assesses the vulnerability of the theater maintenance system. However, our efforts will be concentrated on adding the vulnerability dimension to an existing maintenance simulation — the Maintenance Support Concepts Model (MASC). In brief terms the MASC simulation can be described as follows:

- A queueing model that keeps account of the step-by-step time required for an item to pass through the repair process and be returned to its owner.
- Capable of handling multichannel, multilevel maintenance hierarchies.

The schematic shown in Slide 19-7 illustrates the various activities simulated in MASC. Slide 19-8 illustrates the various actions that are carried out in the preprocessor, main model, and post-processor.

SLIDE 19-7 ACTIVITIES SIMULATED BY MASC



Slide 19-8 — MASC MODEL

Preprocessor

- Generate failure stream
- Generate failure attributes

Stochastic

Main model

- Processes failure through maintenance
- Collect process times from queues
- Add to failure attribute list

Deterministic

Postprocessor

- Aggregate data in history file
- Convert to form of M.O.E.

W.C. Fields' famous statement — or was it Will Rogers? — "There is no such thing as a free lunch!" must have been aimed at models. Recall Pete the Parrot and the requirement for data. MASC is a beast which devours data with gusto. The inputs listed on the left side of Slide 19-9 are difficult to come by. That is a common problem in logistics simulations.

Slide 19-9 — MASC MODEL INPUTS AND OUTPUTS

Maintenance structure and scenario inputs

- Levels of repair
- Number of units at each level
- Combat intensity by unit
- Evacuation scheme for failed items

Policy variable inputs

- Repair facility resources
- Level of repair policies
- Float size and distribution
- TI/QC policies

Probability value inputs

- Correct diagnosis
- Correct repair
- Repair part availability

Distribution parameter inputs

- Failure rates
- Repair times
- Waiting for parts times
- Travel times
- Diagnosis times

MASC Model

Operational availability history

- Overall
- By representative end item

Number of failures generated

- Overall
- By end item
- By end item and failure type
- By final status (repaired or not)
- By transportation type required

Times (Mean and variance by end item and failure type)

- Turnaround time
- Time spent waiting for parts
- Time spent waiting for repair
- Facilities
- Time spent waiting for transportation
- Time spent in repair
- Time spent in diagnosis

In summary, I have described the theater level models that are being used by the TRADOC Logistics Center in studying the logistics system. The paramount thought that I want to leave with you concerns "The Free Lunch" — there ain't none! Logistics simulations require data, data, and more data. And, 'lest we forget, GIGO.

Question: What provisions have you made for getting certain elements of supply like POL from civilian resources in the times of war? Is that a conscious plan?

Mr. Hurford: You're talking civilian resources with the NATO countries?

Question: Yes.

Mr. Hurford: That would be a part of the plan. The answer to that question is "yes." For instance, I had run around making the statement, just to try to get attention, that we don't really have any ammunition problem because we'll run out of POL much sooner than we'll run out of ammunition. We don't really have an ammo problem now, but then there are plans, and Wilbur Payne used to jump on me and say "Hurford you're exaggerating, because we do have in the NATO plan this idea of the availability of service stations as the initial effort and so on — sucking them all dry." That would be a part of the overall system, yes.

Now, let me also say that is going to be one of the difficult arenas as we try to address modeling of the theater level because we now have these host nation agreements, and to try to put them in concrete is going to be a little difficult, because as Joe Heiser says — that's Uncle Joe Heiser, the former DCSLOG who's been doing this study in Europe — unless we have some way to streamline the approaches to that, and this is not just a problem intended for our friends because it's our problem too, unless we can streamline that we'd better set another chair at the table that's doing all these deliberations for the bad guy. That's a real hard part, and we're going to try to address it.

Question: What about other elements of supply, like certain automotive supplies which are common to both civilian and military vehicles, is a study being made to find out what percentage of those you can use?

Mr. Hurford: Well, some of those calls that I was getting yesterday involved the fact that we're doing a high visibility study right now called Tactical Vehical Analysis, and we are being told that we will address, as part of the overall tactical vehicle fleet, NATO commercial vehicles, the M.A.N. truck, the M.A.N. ten-ton, the M.A.N. four-ton and so on. Now, certainly, if we go to the M.A.N. truck, which is a German truck, that is going to complicate our repair part supply problem if we don't have good close-in support out of the commercial industry in Germany. The other problem is we're suppose to do a two and a half-ton truck side analysis. Well, whether you like it or not, a two and a half-ton truck is an ash and trash truck the way we use it, and how the devil we're going to settle by scenario situations that the laundry truck, the kitchen truck, the latrine screen carrying truck, and so on are affected by attack/defense delay I don't know. But the other thing is, they said look at the NATO commercial vehicles, and someone gave us a list of twelve NATO trucks that are competitors for the two and a half-ton trucks in our regular units. But what was not clear was where would we get the data base. We were going to go to DARCOM, but they can't even get good cost data on a few simple vehicles. It's a tough area, and the more we move into that area the more complicated becomes our problem logistically and the more complicated the model.

Question: Why should it become more complicated? Just getting the parts from another vendor?

Mr. Hurford: That's right. We had some good results out of the big transportation unit in Germany, which used the International Harvester trucks. We were supported by International Harvester. Then those trucks got old so we wanted some new ones. The only problem was we wanted them replaced by International Harvester but another guy won the low bid. So, all of a sudden, all of these spares that we'd got at the organizational level, at the direct support level, and so on weren't necessarily going to fit. I don't know whether it was a GMC or what, but the spares wouldn't necessarily fit. It sounds good, as you say, if we can go to a different vendor. But right on one of my slides it said something about waiting for repair parts.

Question: But the point was the M.A.N. vendor is a lot closer than the International Harvester.

Mr. Hurford: Oh yes, agreed. But we also have the 820 series truck, which is a commercial three-quarter ton truck from Chrysler. So we went to that support, and it's working very well in CONUS, but working less well in Germany, which is no surprise, I guess. They want us to support it through the conventional system. You're right, if we could go to the local vendor we'd solve a lot of problems, but then where do we stop using the commercial vehicles? General Vincent said, "No way do I want to see this commercial vehicle you're talking about substitute for the quarter-ton truck in the division area." Well, unfortunately, these problems exist in the modeling arena, and we draw nice blocks on a chart. You know, the first thing you know we've got a line and we've got XX, but unfortunately the logistics system doesn't quite operate that way. You can't have all your commercial vehicles way back here and all your military standards forward. But I'm not trying to take a negative view; we've got to pursue it and maybe we can do it, but it complicates our research. Other questions?

Mr. Spaulding: How do you propose that we ought to look at the interaction between the maintenance

operation and the demand stream? In some ways the demand stream should be influenced by how well the maintenance support is doing.

Mr. Hurford: I wish we could, Stan (Spaulding). In other words, that's the next generation, but we do it pretty much after the fact. As you well know, they fight their battles at Leavenworth and we tell them after the fact whether in fact they could have fought that battle and whether or not they got the results they wanted.

Mr. Spaulding: I guess some cases you get are the worst case demand on the logistics system, where you assume 100% support because if you're not supporting so well your demand stream may be less, but that's a questionable assumption.

Mr. Hurford: Oh yes, we say that all along. It's very questionable. But I don't know, Stan, you've worked in both of these arenas. I don't know how well we're going to be able to interface these kind of logistics models we're talking about with their clutter. They have one heck of a lot of clutter when interfaced with the combat models. If we can get good demand generation out of the combat model, and if we could get good definition of combat damage it would be a different story perhaps. But then the Leavenworth models for the most part tell us how many tanks are hit and then we have to go through the mish-mash with the ordnance school and we have to say, "Well, some of these are catastrophic failures, some are actually reparable at the direct support level, and some require evacuation at depot." We're trying to do that in a sensible analytical way. I wish there were ways that would make the analysis easier.

Col. DuPuy: Have you people at the Logistics Center thought about looking at German World War II data? They had exactly the same kinds of problems you're talking about.

Mr. Hurford: I don't know to what extent we have analyzed it in any depth. I heard your pitch about moving up the Italian boot and so on, and there were some very good insights there. I guess the honest answer is, I'm not aware of the fact that we've looked at that in any depth.

Col. DuPuy: The Germans had all of the problems that you're talking about and surmounted them rather remarkably well in that campaign.

Mr. Hurford: Well, we should look into it.

Dr. Farrell: Our next speaker is Mr. Norig Asbed, from CAA, who's going to talk about a simpler aggregated theater-level combat model and some comparisons that he has done with IDAGAM.

20 — Comparison of Results from IDAGAM with an Aggregated Combat Model

MR. NORIG ASBED
U.S. Army Concepts Analysis Agency

Description of an analytical campaign model and a comparison of model results with IDAGAM.

Mr. Asbed: I'd like to make two observations first before I start. One of them is that I'm going to depart from the traditional pattern of talking about generalities. I'm going to be very specific today, and I'm going really to give you an entire model that you can use if you wish to. The second observation is that, like Dr. Farrell, I too am standing here as a surrogate for Dr. Payne because the model that I'm going to speak about was originated by him, or originally suggested by him. We made some changes in it and we gave it another name, but still it's his model, and, therefore, I wish he had been here to present it himself.

A year ago we had a study that was supposed to be looking into the Army requirements for close air support. We were half way through the study when one day when Dr. Payne was in Washington attending a certain meeting our director was also attending, they had a conversation and Dr. Payne suggested that there was a simple formulation that they could look into if they desired. They then got together on it and later our

director suggested that we try to use it. So we started using it, and as we proceeded we came to the conclusion that we had to make a few improvements or changes to this simple formulation in order to adapt it to our problem.

Apparently Dr. Payne had first suggested this formulation when he was a member of the NATO Research Group I, and the model that evolved was called the STATIC ANALYSIS. This group was actually tasked to determine the forward deployed force requirements on the NATO side in order to blunt a potential break-through attempt by the Warsaw Pact forces in the Central European environment.

Now, the formulation was used again by the NATO Research Group 6 later on. As the name implies, it is a static model in the sense that it takes the requirements, or, it determines the requirements at the beginning of the battle. However, when we looked into it, we had to change it in the way it treated time. Although in our formulation the parameters are not given as the continuous function of time, time is introduced when we determine a first battle and intermediate periods and second battle and so on. But, more importantly, we introduced another fact; we required that the so-called break-through force ratio, or the threshold force ratio, be conserved in the physical sense, namely that that force ratio does not diminish throughout the battles — any one of them.

Now, the importance of this will be shown a little later, and also how we introduced it in the formulation will be shown. I'm going to present the subject in these steps. First I'll talk about the Payne formulation, derive the formula itself, give the scenario, and then from there I'll proceed to the quasi-dynamic formulation where our modifications will be explained. Then I'll discuss some of the parameters that are used in the formulation, how they are determined and their probable ranges, and then I'll go into the application of the quasi-dynamic formulation and compare it with the IDAGAM I, which has been variously described as the outstanding IDAGAM I. Finally, I'll make concluding remarks.

The assumptions underlying the quasi-dynamic or the static analyses are that it will deal with a conventional war, and that the attacker would use the principles of break-through warfare; namely, to amass his forces at the places where he wants to successfully break through and, by the economy of forces, to leave just enough forces in the rest of the corps area, or corps front, so that the enemy will not be able to either counter attack or redeploy his forces in order to strengthen the threatened subsector.

Again, as are most of the models, this is a force ratio model, and again, for simplicity, we are going to assume that in the calculation of the force ratio the attacker's side will only consider the tanks and the APCs, and the defender's side will only consider the tanks and the antitank weapon systems. I should also add that at this time all of these weapon systems have been equally weighted.

Prof. Taylor: So, it's a number count and not a firepower score?

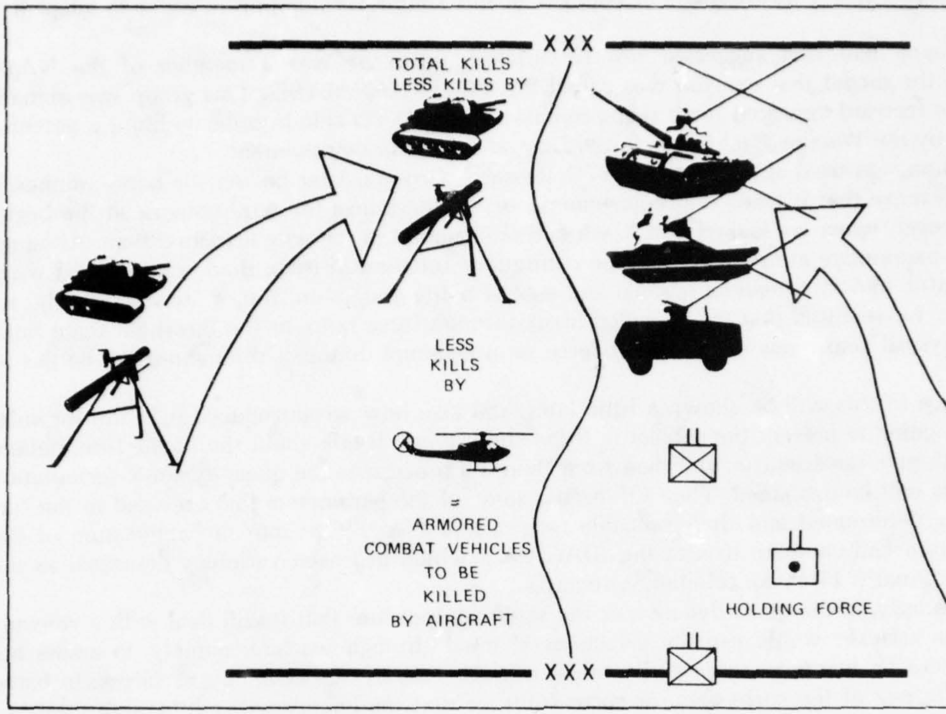
Mr. Asbed: That is correct, absolutely. We make the further assumption that, given the NATO multinational structure, the reserves are limited only to intrasector movement and also that the defender is unaware of what the attacker intends to do, since this was a very short warning war that we were looking into, and will have to uniformly distribute these forces along the entire FEBA to defend, hopefully successfully. All of these assumptions are summarized in Slide 20-1.

Slide 20-1 — ASSUMPTIONS

- Conventional war
- Attacker uses principles of:
  — Massing of forces
  — Economy of force
- Force ratio need only consider:
  — Attacker's tanks and APCs
  — Defender's tanks and AT weapons
- Defender's ground reserves restricted to intrasector movement
- Defender's forces are uniformly distributed along the FEBA

Now, the scenario that was conceived is shown in this picture (Slide 20-2). I have put up here, perhaps a little prematurely, the quasi-dynamic analysis, because this was the original analysis — the scenario — that was envisioned by Dr. Payne. If you look at this you see we have the attacker in the corps area moving some of his forces to a small portion of the corps front, in order to mount a break-through attack, and then he will

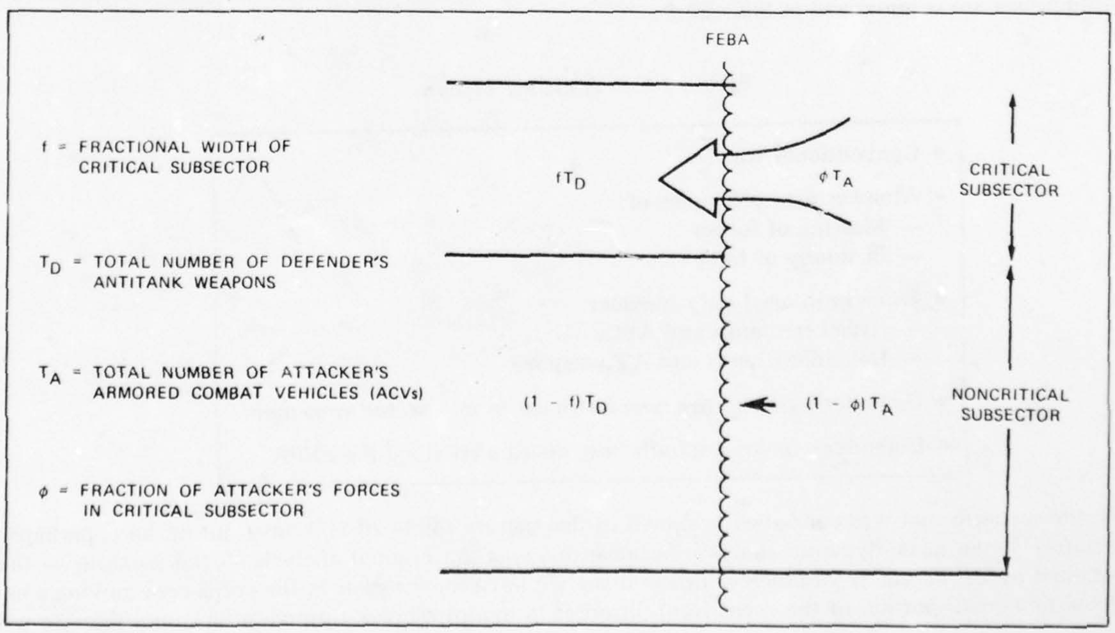
SLIDE 20-2 QUASI-DYNAMIC ANALYSIS



trap a small force that is going to face him on the other side, and he will need only a small holding force in the rest of the corps. Well, the defender, on the other hand, on being apprised of the situation, will move some of his reserve forces only to the critical subsector and then he will calculate his close air support requirements by taking his kills to be derived from air power, and from that he will subtract the kills that the attack helicopters are going to contribute, and then in that way he will determine how much close air support he needs, assuming that he can sustain a certain force ratio, which we can call the threshold force ratio.

The force configuration that I just explained can be formulated as follows (Slide 20-3): the attacking commander will determine a small fraction of the front area, the small fraction, f, of the front area that he

SLIDE 20-3 BREAKTHROUGH FORCE CONFIGURATION



wants to use as a break-through section, and then he will transport a $\emptyset$ proportion of his total attack forces, to that front, and leave behind, of course, 1-$\emptyset$. By doing that, he will have trapped $f \times T_D$ of the total forces of the defender in that critical subsector then the rest of the defender's forces $(1 - f) \times T_D$ will be left in the so-called noncritical subsector.

Now then, taking the force ratios in the two subsectors, the attacker's scheme will give him the force ratios as we find there in Slide 20-4. We have the noncritical initial force ratio that was created in the noncritical sector, given by the first formula on the left, then the initial force ratio in the critical subsector, which is given as that on the right.

Slide 20-4 — BREAKTHROUGH FORMULATION (STATIC ANALYSIS)

- Attacker's scheme

$$R_{NCI} = \frac{(1 - \emptyset)T_A}{(1 - f)T_D} \qquad R_{NCI} = \frac{\emptyset T_A}{fT_D}$$

- Defender's response

$$R_{NC2} = \frac{(1 - \emptyset)T_A}{(1 - f)T_D - k(1 - f)T'_D} \qquad R_{C2} = \frac{\emptyset T_A}{fT_D + k(1 - f)T'_D}$$

k = Fraction of defender's forces in reserve

T'_D = Tank component of defender's antitank weapons

Now in response, the defender has only one course of action at his disposal, and that is he can only move part or all of the reserves that he has in the noncritical subsector from that sector to the critical subsector because he needs some forces to be left over in the noncritical area for its defense. In this case we assume that he's going to move his most mobile and effective anti-tank forces, namely the tanks*.

Now, this, of course, presumes that both attacker and defender are aware of each other's total forces on either side of the FEBA line, and this is not uncommon; computer models will even do that. Now, when the defending commander looks at the force ratio that he is forced to have, and if he finds that force ratio to be very high, in the sense that his forces could not possibly stand it, then he will have to ask for some outside help from other sources. The only thing outside help can do is cut down on the attacker's forces in the numerator.

If we take the defender's response ratio in the critical sector as shown in Slide 21-4 and subtract from the numerator the number of kills that are required from an outside source (close air support) to reduce this ratio to an acceptable level that will blunt the enemy attack, we have defined a new ratio that I will call the threshold force ratio in the critical sector, R_{CTH}. With a little algebraic manipulation, we come up with the two basic formulae that govern the so-called static analysis; one for $\emptyset$ and the other for close air support requirements (Slide 20-5). Now, we notice that in $\emptyset$ there are two quantities that are of interest, the f and the

Slide 20-5 — EXOGENOUS SUPPORT REQUIREMENTS (STATIC ANALYSIS)

$$\emptyset = 1 - \frac{T_D}{T_A} \cdot (I - f) \cdot R_{NCI}$$

$$\Delta CAS = \emptyset T_A - R_{CTH} \{ fT_D + k(I - f)\, T'_D \}$$

R_{CTH} = threshold force ratio in critical sector that ground sector defender can accept and blunt breakthrough operation

*The number of tanks in the defender's antitank force, I've designated as T'_D and k is fraction of the defender's force held in reserve. Therefore in Slide 20-4, $k(1 - f)T'_D$ represents the reserve forces moved from the subcritical sector to the critical sector.

initial force ratio in the noncritical subsector. In the task requirements formulation, which is the defender's formula for his needs, we have this $R_{C.TH}$, which is the threshold force ration in the critical subsector.

The prescription for the subsequent battles which are allowed says that all you have to do is to repeat these formulae, but certainly you'll have to take out the losses and add in reinforcements, and when you do take out the losses you'll have to assume a certain acceptable force losses on the defender's side and a corresponding set of losses on the attacker's side, which is given by the loss exchange ratio.

It is exactly here that the main need for revision of this formulation really comes to the fore, because if you assume that there was a certain first battle and you go into a second battle, and assume a certain percentage force loss that the defender has taken, given a corresponding loss to the attacker, then you can calculate the remaining force ratio at the "end" of the battle. Remember, that this does not determine the length of the battle, it simply says a battle of indefinite time. It does not specify the length of the battle. If one does that then one comes to this formula (Slide 20-6), which determines the new force ratio, R', where p, as I indicated, is the percentage force loss that the defending commander is willing to take, and e is the loss exchange ratio to calculate the losses for the attacker.

Slide 20-6 — FORCE RATIO AT END OF BATTLE (STATIC ANALYSIS)

$$R'(1 - p) = R_{C.TH}\left(1 - p\frac{e}{R_{C.TH}}\right)$$

R' = New Force Ratio

$$\text{where: } p = \frac{k_D}{fT_D + k_D(1 - f)\,T'_D};$$

$$e = \frac{k_A}{k_D} = \frac{k_A}{p\,\{fT_D + k(1 - f)\,T'_D\}}$$

and k_A = number of attackers killed

k_D = number of defenders killed

This formula clearly tells us that unless the loss exchange ratio is equal to or larger than the critical threshold force ratio, the new force ratio is going to be larger than the critical threshold force ratio. Now, this simply tells us that the break-through attempt has succeeded, and our attempt at stopping it has failed, because we postulated this threshold force ratio as the one required, or as the one perceived by the defending commander that he can take and yet blunt the break-through. He has not succeeded in that and there is no reason why the attacker should stop unless one brings in some extraneous reasons such as logistics and fatigue, or some such other factor.

Well, when we looked at trying to resolve this problem, we had one other aid in the sense that we had attack helicopters in our problem, and the data for attack helicopters was such that if you wanted to calculate the attack helicopter kills you had to have the sorties per day of attack helicopters, their availability, their kills per sortie, etc. (Slide 20-7). Therefore we had to estimate how long the first battle was going to last, how

Slide 20-7 — ATTACK HELICOPTER CONTRIBUTION

$$\delta_{AH} = N_{AH} \cdot AV \cdot SPD \cdot KPS$$

$$L_{AH} = N_{AH} \cdot AV \cdot SPD \cdot LPS$$

δ_{AH} = Daily number of attack helicopter kills

N_{AH} = Daily number of attack helicopters

AV = Availability of attack helicopters

SPD = Sorties per day per available attack helicopter

KPS = Kills per sortie

L_{AH} = Daily losses of attack helicopters

LPS = Losses per sortie

long there was an interim period, and how long a second battle could last.

Well, our battle, the duration of our war, was fifteen days, and the reinforcements were going to arrive and be on the front after the third day; that is on the fourth day they would be on the front. Therefore, we postulated that the first battle had to last for three days, and so we made our calculations of the attack helicopter contributions to the kill on that basis. But now, more importantly, we had to make sure that the threshold force ratio remained constant, and this we did by demanding that the defending commander preserve, or conserve, his forces. To ensure this, we take the threshold force ratio always to be with the remainder of the forces rather than with the forces at the beginning of the battle.

Well, on doing that and rewriting the equations, we come to the two equations of the quasi-dynamic analysis (Slide 20-8). The $\emptyset$ does not change from Slide 20-5; that is the attacking commander's contribution, and that is his calculation. We cannot do much about that. But now the defending commander has introduced some parameters. Let's start with the bracketed expression first. Notice that the first two terms of this expression were the first two terms that we had before, but now we have introduced another term on the basis that it is known Russian doctrine that they will mount a severe preparatory firing at the beginning of each battle, especially a break-through battle. Therefore we assume that they're going to direct all of their fire against the antitank weapon systems, and that's where you get T_D-T'_D and T_D. The prime refers to the tanks and T refers to the total; so, if you take the tanks from the total you're left with the antitank weapons since we assume that the defender's forces are only made up of tanks and antitank weapon systems.

Slide 20-8 — EXOGENOUS SUPPORT REQUIREMENTS (QUASI-DYNAMIC ANALYSIS)

$$\emptyset = 1 - \frac{T_D}{T_A}(1\text{-}f)R_{NCI}$$

$$\Delta\,CAS = \emptyset T_A - R_{C,TH}\{\ \} - pe\{\ \} + p\,R_{C,TH}\{\ \} - \delta_{AH}$$

$$\{\ \} = fT_D + k(1 - f)T'_D - Ef(T_D - T'_D)$$

p = Defender's acceptable proportional loss

e = Loss exchange ratio (attacker/defender)

δ_{AH} = Helicopter kills of ACVs

E = Fraction of non-tank antitank weapons lost to preparatory firing

Here, again, is the other important element that I would like to point out: that is the next to last factor in the expression for ΔCAS. This factor reflects the losses of the defender, and we see that the loss factor is added to the close air support task requirements, which means that the close air support has to take up the slack left from the losses of the ground forces. By the same token, though, the defender has been helped in his close air support requirements by the amount of kills that the ground forces inflict on the enemy forces. Finally, we have the attack helicopter contribution, and we calculate this according to a very simple formula, which simply takes the product of the number of helicopters, their availability and sorties per day, times the kills per sorties. But now we also have to calculate the losses at the end of every day in order to have the remainder of the helicopters for the next day, and these losses are calculated by a similar equation (Slide 20-7).

Let me put Slide 20-3 back on and look at some of the parameters and their importance. First, f is the parameter that the attacking commander determines at the beginning. He is the corps commander, and he will decide what fraction of the corps front he is going to use for his break-through. Now, in this decision are a few factors that he has to look into. They could be in the form of what is the optimum force that he would like to concentrate in a given area and whether he could concentrate them in a given area. You can only put tanks, say, within 50 m of each other, or something like that. He will have to consider the demography and otherwise the geographic locations of his objectives, whether they fall within that break-through area, and certainly he will have to consider the terrain and other natural barriers that fall in this area.

Now, the accepted values for a corps front for a breakthrough are between 10 and 20 km. We have assumed a value of 15 km so f is always given to be such that it will reduce the break-through front of the corps front to 15 km only.

The next thing that the attacking commander considers is the initial force ratio in the noncritical sector.

Now, this one is a translation of his economy of force principle into the formulation because he, by that number, determines how much force he can optimally leave behind and yet be able to mount a strong break-through attempt. What he leaves behind is going to serve several purposes. He may try to prevent a counter attack, he may force a stalemate, or he may frustrate a redeployment of forces into the critical subsector. Now, for this one, numbers anywhere from 1 to 3 have been used. We have used 1.5.

Then comes the critical threshold force ratio. Now, this one is determined by the defending commander, and here he has to look into the morale of his defenders, he has to look into the relative technical skills of the two forces, or his estimates of them, and he has to look into the relative capabilities and sophistications of the weapon systems at his command. I should note that this critical threshold force ratio can be selected to accommodate several of the simplifying assumptions that have been made. For instance, equal weighting of the weapons can be absorbed in this as can the inclusion of only certain types of the weapons but not all of them. Those of you who have used IDAGAM will immediately point out to me that IDAGAM includes all of the weapon systems in the calculation of the force ratio except airplanes, but here we're including only two, so the value that you use for this can be adjusted for all of those differences with other methods of force ratio models.

The p, which is the acceptable loss to the defending commander, is one that has to be looked into again. It has two considerations: One, he has to make sure that his forces are always combat-effective; therefore you have a maximum of 39%, or 40% according to some of our instructions — whether this is generally acceptable or not I'm in no position to judge; two, if he is to conserve p, he has to stay mobile. If the attack is too strong he has to move away — he has to disengage. We have allowed p to be 25%.

Now, when we tried to apply these to our problem in determining the task requirements and compare them with IDAGAM results, we considered that in each corps area there was going to be a similar attempt made by the enemy corps commander to create a break-through area and noncritical sector.

In most of the sectors, or the corps areas, there was no close air support requirement because the respective forces did not produce a large enough force ratio. Incidentally, we took the critical force ratio to be 4. We then aggregated all of the corps area results and compared these with IDAGAM (Slide 20-9). We looked into three types of force structures for 1977 and 1982. The first one is for 1977, and the other two are in the 1982 time period. The task requirements given as percentages of IDAGAM are 99, 93 and 115%. The comparison is strikingly close. The 15% discrepancy in the case of Force Structure 3 has an explanation, and that is that we had decided when we were doing the quasi-dynamic analysis that we were going to select the parameters only once, we were not going to fiddle with them after we got the IDAGAM results, so we left them as they were. However, there was an element that our model was equipped to handle, but we did not use it in our calculations; namely, in the third case there was a weapon system improvement. Such improvements were included in the IDAGAM runs, but in our case we did not include them, and therefore our requirements were much higher than the IDAGAM requirements.

Slide 20-9 — IDAGAM-1-QD RESULTS

Force Structure	CAS Requirement (% of IDAGAM-1)
1	99
2	93
3	115

I should also point out a few other differences. One of them is that IDAGAM 1 fought a 15-day battle, and a continuous battle; it accumulated all of the close air support requirements in order to do that. However, to find the close air support requirement, the IDAGAM model demanded that the FEBA not move, or at least remain very close to the initial FEBA.

Now, we have assumed that p, which is the loss that the defending commander is willing to take, is of the order of 25%. Now, it would be highly unlikely that the defending commander could stop at a FEBA that remains fixed and fight and only lose 25% of his forces. Therefore, the quasi-dynamic calculations will certainly include a FEBA movement. It does not predict what it is, but it does require that the forces move, especially if you're demanding that your losses be no more than 25%.

The other important point is that in the quasi-dynamic model, the essential requirements for close air support were concentrated only in the first three and the last three days of the war. In between, it was only a

5% loss to the enemy force, which was very minimal. Now, when your requirements are concentrated only in time and space with such limits this will put constraints on your capabilities. Therefore, this probably will suggest some requirements that the Air Force and TAC/AIR will have to satisfy.

In conclusion (Slide 20-10) I'd just like to say that it's a very simple model, it's extremely transparent, it's very responsive, and you can do the whole calculation at relatively minimal cost. I was going to say at no cost first, but then my director objected. He thought he had paid me for doing the calculations. Then, the model really gives realistic surge requirements because the requirements are concentrated in time and space. It does not, however, predict FEBA movement.

Slide 20-10 — CHARACTERISTICS OF THE QD MODEL

- Simple, transparent, responsive — relatively minimal cost
- Realistic surge requirements from CAS
- Does not predict FEBA movement

As a final statement, I would like to say that if simplicity and responsiveness are really virtues of modeling, then we have found the quasi-dynamic model very useful as a tool, or as a back-up to some other analyses, some other analyses that are being used especially for force and force structure requirements.

Dr. Farrell: We have one more speaker for Session III and then we'll have the panel. The speaker will be Robbie Robinson from Air Force Studies & Analysis, talking about tactical decision processes in theater models.

21 — Modeling of Tactical Decision Processes in Theater-Level Gaming

MR. ROBERT ROBINSON
HQ, USAF (AFSA)

State of the art in command and control in theater models — recommendations for future research.

Mr. Robinson: I'm going to try to do something that appears to be somewhat of a novelty and address the subject that I was asked to talk about.

The subject that I was asked to talk about is an overview of the methods of decision making in theater-level gaming. I have a series of slides here that outline the overview.

Briefly, in terms of the history, war games have been around for a long time, and they've been used successfully, and should continue to be so used, in the training of people and the testing of procedures and the development of operational concepts to gain an understanding of the nature of combat and what goes on there. Through the years, of course, they have required computer assistance, but with man still strongly interacting within those games. However, if we are interested in the study of the decision process, it is fairly hard to track what those decision processes were when all of the decision occurs in the minds of the players.

When we play closed games with umpires and all of that, at least we have some track record, but basically the decision process and the perceptions of the situation occur within the mind of the players, so if we have any reason to try to understand the decision process, we can learn something through the players. Although there is a terribly poor record of what they did and why they did it, we can achieve a lot of understanding.

Now, there are a lot of these closed games around, JIFFY and all of these that have gone on through the years, and I don't expect them to go away because they are certainly very useful. They were especially useful in the period of analysis when we started using theater-level simulations that ranged all the way from the TAGS model of the early 1950s up through today's ATLAS and IDAGAM. These models have a particular characteristic that I can identify with those simulations in that most of the representation of the decision process within those simulations is at the theater level and the decisions that are made are largely decisions to allocate reserves or air resources or supplies or something to a particular sector and from then on below that there is little in the way of a decision process. When we get out to the front line the forces are always moving forward vigorously and staying in there and fighting regardless of how long or how serious or intense the engagements.

Presumably, there is some decision process that affects those things, and if we are going to understand those combat processes as well as the decision process, perhaps we should reflect something of the nature of those events.

As it was observed yesterday, one of the purposes of this conference is to consider the role of game theoretic approaches to modeling, decision processes, and so forth. In the early 1970s, a number of models were developed; TAC CONTENDER among them, the pride and joy of AFSA for awhile, and from which a number of interesting things were learned. However, no matter how the process is represented or calculated, these models assume an extraordinarily logical commander who has perfect information available to him and who makes decisions on a very precise schedule. Moreover, the models of combat that are associated with those models are generally very simple because of the computational difficulties, length of computation, and all the rest that goes with the game theoretic approaches, so there are limitations there too, even though we have learned a great deal from those particular developments.

I've already raised briefly the question of why should we be interested in modeling the decision process (Slide 21-1). One of them is just for pure credibility. I think a number of us here have been berated with great passion by Tony Cordesman in the past about the importance of putting models of command and control in our analyses just to represent the reality of the world, to give it some credibility. Andy Marshall mentioned the same sort of approach yesterday, although I must admit that later in the day General Kent said that we should never model the decision process. So I'm caught somewhere in between those perspectives. I believe that if our simulation of campaigns is to be credible we need to have some representation of the command and control system. That leaves a lot of room for deciding how much representation. I would also contend that, if we are going to properly do this job of assessing strategies and tactics and weapon systems and force structures, the command and control system should have some representation, because I think there have been a good many observations made that, as we change the weapon systems and the mixed systems within the force structure, the tactics and the strategies that are used should also change. I think this sort of thing is evident from the game theoretic models that were developed and used in the earlier 1970s. And, of course, finally, I'm interested in assessing the value of C^3 systems themselves, and this hasn't been done in the past very well, if at all.

Slide 21-1 — PURPOSE OF DECISION MODELING

- Credibility
- Value of C^3I
- Strategy, tactics, force structure

What is the value of spending more on information rather than buying additional weapon systems? This is not well addressed, let alone well answered. I think Jerry (Bracken) we're probably a reasonable way off from being able to satisfactorily handle the question of trade-offs between tanks and aircraft to be able in a single analysis with a single construct hierarchy or super model, to do the problem of information versus a weapon. I don't think we have a handle on that one at all, although we are making slow progress in this direction.

Now, I think there's been a fair amount of discussion this morning and yesterday, too, that at least we have some theories for modeling the combat interactions. I was going to use an example of the Lanchester theory, but I'm a little afraid of that now after hearing the variety of the discussion, but at least theories exist. But as a model or a theory of command and control, I don't think that exists. There's still a significant gap

that must be bridged between the decision theory as it is talked and developed, and a theory of a command and control or a management system. So, that is one of the major bridges that we must cross.

Now, I've talked about command and control, and I think that before I go any further I should make a couple of definitions so that at least I can say that what I was really talking about was this, and not your definition.

I'm defining command as being the decision and planning process and the action that comes out of that is the issuance of a plan and some orders of direction of guidance. The control part is the attempt to make that plan come true. The endeavor is to produce a coherence in the force to maintain it as a stable operation. Plans may change and the orders that come up from different places may conflict, but the total control system is to try to bring some order to the house.

Also, I have been using the term information as contrasted to intelligence. I consider intelligence in the usual sense of a subset of information, and the information that I'm talking about is the total set of information that becomes intelligence about the enemy forces as well as the status of one's own forces and other characteristics, ideas, and notions that one should have about one's own operations.

Now, I will propose that there are a number of essential elements to command control communications and information (Slide 21-2). I won't claim that these are all the possible elements, but I suggest or postulate that these things at least must be contained, that in my command process, my representation of the command staff, I must include an information input, and then I must represent the process by which the situation is perceived. I think that is very important because the information that I put into this system is incomplete, it's erroneous, and frequently out of date so the picture that the commander and his staff has is a very fuzzy picture and from that he must perceive some situation and then progress to an anticipation of what the future may bring him from that particular situation. Then, he will make some plans that will either endeavor to cause that anticipation to happen or not to happen depending on his particular goals and the attrition that he is suffering, the rates of which events are happening, and so forth. Finally, out of that comes an action, which, as I mentioned before, is in terms of orders, directions, and guidance for the control system to take over.

Slide 21-2 — ESSENTIAL ELEMENTS OF C^3I

- Decision making
  - Information input
  - Situation perception
  - Anticipated situation
  - Planning/decision
  - Action
- Hierarchies (air/ground)
  - Space horizon
  - Time horizon
- Uncertainties
  - Events
  - Information
  - Communications
  - Decisions

Now, another fundamental element of modeling the command/decision process is hierarchy. I think that if you look in military textbooks as well as analysts textbooks you will find hierarchy mentioned, but what is not treated are the hierarchy elements — the levels in the hierarchy — and these have important differences in terms of the space and time horizon. What do I mean by that? I mean that the highest level of command, the planning horizon in terms of the space viewed by the corps commander covers that total corps area, or the theater commander covers the total theater, and the time space in which he must make a decision is a longer one. His planning horizon, his time planning horizon is greater, while at the engaged forces level the time horizon is very short and the focus of interest is much less, the area in which he is interested is much less.

Finally, there are a number of uncertainties. If we could treat this total process of command decision making and control deterministically, the problem would be greatly simplified, although we haven't worked it

very well yet, with a few exceptions. First, there are uncertainties in the events that occur; in spite of a force ratio of 4:1, the attacker may lose. It might not lead to a break-through. It can still lead to a draw because of the uncertainties of the events that can occur. There also are uncertainties in the arrival of supplies as well as uncertainties in the occurrence of combat interactions, and all the way on up through the tree.

Second, there are a lot of uncertainties about information of the kind that I mentioned previously. It may be in error; it may have other limits; it may not arrive in time; on and on. The information can be completely false and the commander may perceive his situation based on that falsehood.

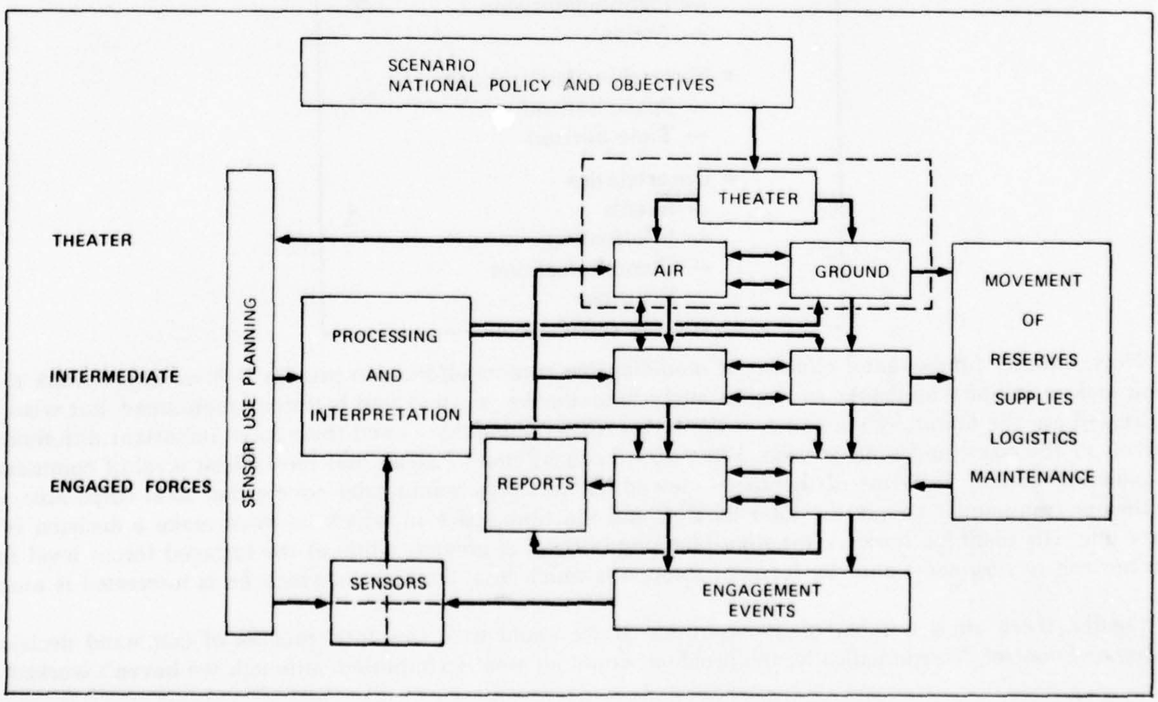
Third, regarding uncertainties in communications, delays in arrival, error rates, and all the rest, I think we generally know how to treat that part of the problem rather well.

Finally, there are uncertainties in decision. Even if the commander follows doctrine, you're still not quite sure what he's going to do because most written doctrine leaves him sufficient latitude to do a number of different things. Also, there is really no reason to believe that, in a given situation, a commander will follow the doctrine. I don't care whose doctrine it is, and there are a variety to choose from. However, the purpose of our modeling is not to measure the characteristics of different real or postulated commanders. The purpose in our modeling, I believe, is to measure the characteristics of the command and control and information system and find out if it can be responsive to a variety of commanders and situations. So that I would propose that among these uncertainties this fourth one (decisions) is perhaps the one that I would treat deterministically — the one variable that I would hang onto and say "Well, I'm going to ignore the normal distributions or whatever else I might have, to represent that commander."

There's something else here to consider. Into this process as I modeled it, I have a number of inputs in the modeling of command and control. My input is information, intelligence, and status of forces, so if I'm going to model command and control I must represent that input system as well. Moreover, you have to remember that the output of the command and control is something that is not substantive like things killed, probabilities of hit, but orders, direction, and guidance. And there are certainly uncertainties in those things as well.

Now, to tie all these things together, I'll use a simplified picture (Slide 21-3). I don't have both sides on this picture, and the structure that I show is essentially that of the U.S. Army and Air Force, but it need not be that way. I think one can generalize from this sort of thing, to Russian or German or British, even though I understand that the British are going to leave out brigades and the French are going to leave out divisions. I haven't worried that problem too much, but have just showed layers to get my hierarchy so that I have layers

SLIDE 21-3 HIERARCHY IN COMMAND/DECISION PROCESS



of command. I have engaged forces, I have an intermediate level of command and the theater level, in which I've boxed a number of things together. I have an input to start my game, the scenario and national policy objectives, and so the initial orders would be to go. I think I remember hearing General Welch remark awhile back that one of the differences between the modelers of military operations and the modelers of economic events is that the economist never is concerned about whether his units are going to engage. Well, of course, one way to keep your attrition down is don't go, and this, then, is a problem for a commander — to cause his forces to engage — and that is the sort of thing that must be represented. Of course, the objectives and goals at the different levels of this command structure are also different. The objective of the theater commander of course is this overall one to either win the war in some terms, that is, occupy the amount of territory that he has been directed to occupy, or to prevent someone from occupying that territory. But the guy at the lowest level may have as his first objective "don't get killed," and these objectives may conflict. So, we must clearly identify that there are different levels and there are different objectives at these different levels.

Now, I have represented in a box of Slide 21-3 the logistics and all of those other good things that Mr. Hurford referred to, so I haven't ignored him completely in this discussion. But I have given a much larger role in my picture to the role of the information system. But even that is simplified because I contend that the process of planning for the use of sensors, and the decision making that is made there, may be as important as the process of collecting the information itself. Some recent analyses have shown that sensor use planning may be as important to winning some of the battles as planning for the use of the fire and maneuver forces. You can't do without the fire and maneuver forces, however, some action can be taken without the appearance of new information on the scene. Then, of course, you don't ignore fire and maneuver — we must model the engagement events as well in spite of the fact that there still seems to be room for contention over the validity of the theories of modeling combat interactions.

Now, where are we today (Slide 21-4)? In theater-level models, we've had considerable mention of CEM and VECTOR 1, VECTOR 2 — All of these have command hierarchies in them for ground forces. They do not represent the interactions with the command structures with the Air Force, but at least we've gotten to the point that we recognize there is a command structure and these two models do contain them. CAA is working on the further extension of CEM, adding command control communications, intelligence and electronic warfare, under contract with Vector Research. The model will be largely but not entirely deterministic, and so that means, to me at least, that what we will model there will be according to doctrine and that we will perhaps insert as in the VECTOR models, sets of decision rules. So there is considerable left to the analyst to put in decision rules that reflect the world that they're trying to model.

There is virtually no Air Force, no air operations in the CAA CEM model, even as improved. Since I represent the Air Force, I like to believe that there is some impact caused by those operations, and I do believe that there is a considerable interaction between the command hierarchies of the air and ground forces.

Slide 21-4 — CURRENT MODELS AND APPROACHES

- Theater
- Lower level
- Game theoretic
- Artificial intelligence

At lower levels of models, one of the most interesting I've seen in a long time is the FOURCE which is the grandson of DIVOPS, and a number of other things, which was developed under Dr. Payne's direction. It, too, is a deterministic model. It contains a highly structured set of decision rules, and I think they've done a beautiful job on it for evaluating the Tactical Operation System (TOS). It is, as I said, deterministic, and therefore it is confined to the assessment of that system in the context of doctrine, but that is a major step forward. This sort of thing hasn't been done before, and I think it's an extremely interesting approach. Incidentally, I should mention that FOURCE has no Air Force in it at all. In the Air Force, where we usually leave out the ground, I'll start with TAC AVENGER, which has been around for a long time and is a model of one-on-one engagements, air-to-air duels. It does have a decision process for the pilots in which the characteristics of the aircraft, the perception of his opponent in the duel, and the possibility of maneuvers and energy that he has left will govern his actions.

Under development in AFSA is an air-to-air model, a flight-on-flight, using the decision processes of the kind that Dr. Pugh described yesterday. It is a value-driven model that should give us considerably greater understanding when we get away from the one-on-one. Then there is also in AFSA a model of air-to-ground operations under development, flight on a target array, that also contains guns and surface-to-air missile systems of a variety of kinds. Among the decision processes that would be included within that would be the kind of decision that a pilot would make when he has perceived on his RHAW gear (radar homing and warning) that he is under attack by a SAM. So there are steps that are being made that should give us insights that are applicable to the larger scale model.

Now there have been several honorable mentions made here of TAC EVALUATER. This development was initiated at AFSA and has now been observed at Nellis, and at other points around the globe, not too closely always. It is an entirely deterministic ground battle that is built on the SCORES EUROPE 2A and in its present form is a representation of that particular scenario and that one alone, with small variances allowed around that particular center. It is being used now in an analysis to evaluate aerial reconnaissance systems, and it is also being used as a test bed for the development of C^3I models, command, control, communications and information modeling techniques — a methodology.

In the game theoretic business I've seen very little new in applications in the recent years. I made a suggestion for a possible application in a paper I gave to explore the idea of using the game theoretic model, or technique, at the highest level, the theater level of command. One purpose was to try to overcome some of the objections that are generally made about game theoretic formulations and to suggest that the theater commanders of each side play a different game, — each plays his own game against his own objectives, and with his own planning horizon. Instead of running the game out until the end of the war, you run it to what you estimate that commander's planning horizon to be and that should take some reasonable judgment, which is not impossible.

Also, game theoretic approaches have been critized for the fact that the objectives remain invariant with the war. I think that we can write our payoff equation, and these generally are written in the linear form with constant coefficients. I can see no reason why these coefficients are not variable, with the progress of the war depending on the rate at which objectives are being achieved, and at which attrition has been taken, considering attrition as an absolute or as a rate so that the weighting values for the payoff equation could change with time. This is the same sort of scheme that is involved in using value-driven weighted decision rules and the sort of thing that has been appearing in a number of different areas, such as artificial intelligence. There has been a lot of work done here in recent years and some of it is getting very interesting. Chess-playing computers have improved greatly. I've seen that you can buy them on the airplane now, if you've got enough money or your credit card, and can play a fair-to-middling game. The better chess-playing machines today don't operate on decision trees; they operate more on a pattern recognition concept and don't try to solve through to the end of the total game.

Second, other developments have been occurring in robots with hierarchical decision processes built into them. These robots can do a lot of things other than the bugs that used to be demonstrated and looked so fascinating. AFSA is using an artificial intelligence concept in its research for representing command and control, structuring it out of the idea of pattern recognition. We feel that this idea is key although we're not far enough along and convincing enough in our argument not to be open-minded enough to see other possible approaches.

Now, decision rules are built into all models one way or another, and the decision rules of the simplest models are from open games played by people, which were discussed at the banquet last night. There are decision rules there, rules of the game. We go from these to the kinds of decision rules that were condemned yesterday, if you will, the highly elaborate, fixed, decision-rule process. These, in turn, can lead to some very strange outputs, so that, if we're going to represent the dynamics of the command and control process, there has got to be some variability in this decision-rule process. It has to involve a value system, a weighting system, that changes dynamically according to some events or situations within the operation of the model.

I'm going to stop here. I do have a suggestion for some research (Slide 21-5), and number one is input data. I think there are places that we can get input data that have not been plumbed. These include the better and more careful examination of existing command and control communication systems. We can use our CPX, or, command post exercises, even though one must always be cautious how they are used, and there's a great mine of information, I should hope, in historical data.

I think we need an extended research program using manned interactive simulations that range from the war games of the theater to specific rooms in which people perform specific roles. I think that latter is particularly important to the understanding of command and control and also to the design of better command and control systems. It's a favorite subject of mine, and I mustn't get wrapped up in it. It's a source of data,

and there are experiments that should be done that should lead us to a great deal of better understanding than we presently have.

Finally, we need a theory of command and control, and we need to develop models to work it out, use it, try it, and see if that theory has any meaning.

Slide 21-5 — RESEARCH

- Input data
- Manned simulation
- Theory and model

Panel Discussion — Session III

Mr. Farrell: If the panel will come up, we can have the panel and questions from the audience at the end of the panel for any of the speakers during the session except Mr. Hurford, who has gone.

I'm going to break with tradition here and not take people in alphabetical order. I'm just going to take the people in the order they sat down along the table, and we'll start then with Klaus Niemeyer, of IABG.

Mr. Niemeyer: I'd like to pick up your introductory remarks this morning and give some ideas of what we do in Germany on this kind of model work. I think some of you already know what we do, what type of models we have available, what type of models we have developed, but I'd like to remind you that I'm talking about RELACS, which is a theater-level model. It is an integrated air/land war model. It is used as an interactive war gaming model, and this is more or less the top model of a hierarchy of models which provides the basic data in an aggregated form to this model.

Very closely connected to RELACS, is a model we developed that we call TREND. This is an even higher aggregated model based on RELACS, but uses more or less aggregated decision processes based on simulation runs with our RELACS model.

Basically, we found in using our models that all the problems that have been raised here during the meeting came up in the same way, so I think this is very common.

I think also that what we have learned this morning, and also yesterday, that basically the problems of the attrition of ground, air, and marine systems have been solved more or less adequately.

I also think a good job has been done on logistics, as we have heard this morning. This, I think, is due to the fact that basically the assessment of losses, of rates, and so on is based more or less on physical data and that these data are relatively easy to relate and to understand.

Also, the experience with the MCSSG study showed that, as Rex Goad pointed out, there are not too many differences between the models which were used in the MCSSG comparison. From our side, it was RELACS, and from the United States, of course, it was IDAGAM. Maybe we will have some more comments regarding the British models from Dr. Dare, but I think that lots of problems remain in this field, especially the problem of where the human interaction gets into the analysis, particularly in the assessment of unconventional warfare. We have heard about this problem this morning, as well as yesterday. I think urban warfare is also a very interesting field, which hasn't been analyzed to an extent that we are able to assess the effects of this type of war.

Also, the interactions with civilians, the behavior of civilians in war, hasn't been evaluated to a great extent.

Last, but not least, the decision process is a very important thing, which has to be evaluated. The MCSSG study also shows that here are the most differences in between the models that have been used. I think the last presentation was very interesting in this respect, because we might get some new ideas out of the evaluation which has been done in this area.

In my opinion, the decision process is very important, especially at the higher levels, at the theater level. Here, we have to do it with force deployments, we have to do it with the commitment of research, for example, and we have to do it with the reaction of the defender, reactions to the center of gravity, or the center of attacks of the attacker. The sensitivity analysis we have done in this area with our RELACS model, also with our TREND model, show especially that those center of gravity considerations are very important. Based on this experience, we think we have to focus our attention much more to this point.

Another thing in connection with this decision

process is that in reality the decision maker has to make the decision and, of course, he doesn't like to be represented, let's say, by some formulas, or artificial intelligence on the computer. I think this is very important to understand, and we have to take care of that when we want to explain our findings and models to these decision makers.

Decision making can be split up into two areas. The first area is decision making in peacetime. This, of course, is a very complex problem and is based on the balance of power between all the organizations, services, staffs, and what-have-you. It is not simple, as Frank Kapper pointed out yesterday. We think, or, I think, that this type of decision making has to be analyzed in the same way as we do it with the other activities — with the other systems. That is, we have to find out where the best place is, what the best time is, and what the best formula is to bring the results of our models into this decision making process in order to convince our decision makers that we do a good job. It's my opinion that we have to analyze this decision making process, and second, that we have to try, at least, to get a dialog going between the decision making community and the analysis community.

The second type of decision making is, of course, decision making in wartime. Here, we have the same people. We have no chance of doing any analysis in a wartime situation because we do not have the time to do that, and decision making in this area is only based on the knowledge and the experience of those decision makers. In the future, they won't have any optimization routines and devices available to do optimization in the field, so decisions are really based on what they know, what they have learned, and what they've gotten from insights and experience in peacetime. The only way to get this experience is using models which we've developed in the past.

For this reason, I think the models we have developed are more or less an experimental device to get insights about the phenomena of war and that the people who use it, especially the decision makers, are trained to gain the knowledge to do the job in war.

I think both the dialog between the decision makers and the analysis teams, and also the knowledge and training gained for wartime situations, is necessary. So for this purpose I think interactive war games serve best. Over a period of time, I think the analyst would be able to analyze this decision making process and to eventually model it to develop, in turn, more aggregated models, which then can be used for force structure analysis. Ultimately it would not be necessary to look into this decision process in detail. You could avoid some of the specific aspects of the decision process, and concentrate on getting trade-offs between weapon systems and force structures.

Thus, the combined use of war gaming and closed simulation model, the TREND model, with an aggregated decision process in it, is the best way to get the answers we are looking for. The advantages of this approach are many: (1) the dialog with the decision maker is going on, (2) the training of the decision maker for wartime might be better, or might go on in wartime, (3) these people get insights to the phenomena of war, and (4) these insights are not only accruing in the analyst community, which is doing the model work, but also is getting to the people who really use it and need it. The research work which has been mentioned, for example, in the last presentation can be done, and finally the force structure analysis for force planning can be done when these models, these closed simulation models, are developed based on the experience that comes out of the war games.

Mr. Farrell: I think it will be easier, since we only have this one live mike if we just pass it on down the table and I let people introduce themselves.

Mr. Dunnigan: My name is Jim Dunnigan. I'm a commercial game designer, so to speak. I'll explain what we're about. We design commercial games. I used to do it for Avalon Hill Company. Now I have my own company, which fortunately got most of the market since I did that. There are about a million of these games sold a year. They're primarily highly simplistic versions of what you fellows do, although I've got some people on my staff who would love to commandeer a computer and do nothing but crunch numbers for days on end. I've had to rack my brain to try and find some information, some observations that could be of value to you. There's much that I've picked up here that's of value to us.

Two things I think I've been able to come up with are the use of playability and play testing as means of producing greater realism in the models that you use. Now, what I mean by ease of use, or playability, is something we are always concerned with since we are basically selling our products to a lot of people like yourselves, a lot of people in the military, but primarily interested civilians. Our average customer has a few years of graduate school, is male, from 18 to 34 years old, and rich enough to indulge in our products, which aren't all that expensive. They have enough time and they have a great deal of interest and curiosity, so our products have a lot of information. The information has to be fairly obvious, and it has to be, of course, playable. That's the only way we can gain the customers. If we scare them off, they're gone and we're gone. So, we constantly have to face this problem of playability. I would suggest that when putting together models for institutional users, shall we say, which is what most of you deal with, you pay more attention to the human factors. The fact that in order to get your message across, to have all of the work that you have done in constructing, you know, a valid model, used, it has to be presented in a way that the user will be tempted, shall we say. You can do now what you couldn't do previously. For example, you could stick two CRT terminals in some general's office

or whatever, some decision maker's office or somewhere where he can get hold of them, and you could have perhaps a simplified version of your model. When I say simplified I mean that there are only a few parameters that the user will play with. It will look a lot like some of the games you're going to see coming out for television, the Pong type of game. RCA has one, Fairchild has a somewhat simpler one, and this game technology will be catching up with you whether you adapt to it or not. In effect, you can put the user in control of this powerful model and allow him or her to play with it. Don't underestimate the value of play. You know, I've had letters from people, generals and other high ranking civilian decision makers, who play our games, just as a number of you people play our games.

Most people "play" our games solitarily. They don't use an opponent, even though the games are normally for two players, and this is the thing that hooks them. The idea that they have this pile of information, this movable data base, which is basically what your models are, and to a much simpler extent what our models are, people want to be able to assess it, to be able to play with it. Now, when I say play, I mean quite frankly the fellow might sit down and just sort of stare at it like you'd stare at the 7 o'clock news. He may not understand everything that's coming across, but every once in awhile he'll say, "Well, look at that, I just wiped out half of East Germany" or something like that.

I don't mean to be facetious, but it is important to deal with the human factor, and that seems to be one of the biggest problems you've had, getting the message across. There's obviously a failure of communication so to speak, and from my experience this communication problem can be overcome.

Now, another observation, in terms of model building — how we build models. We build very simplistic models. Practically everything we do is aggregation and preprocessing; we never get beyond that in many respects. In fact, I've stolen a number of good ideas and I'm working on my second page of ideas from you folks. We are doing a game right now, and this is a coincidence, a theater-level model so to speak, of the central front. We did one about three years ago called "NATO." Of course, we can't use that in the next version. So we're coming up with a new one in the next war, very imaginative. This is the interesting aspect of our work. About a third of our games are on current and future topics; the rest of them are either World War II or those that go all the way back to the dawn of history. We can do the current topics, again, every three or four years since everything becomes obsolete, which is nice from a marketing point of view where the topic never becomes obsolete

Every time we do a new game, of course, we try and repair the systemic errors that were in the original one and update the information; we try to make it bigger, better, and more playable. In doing this, we use an approach we call recreation as opposed to calculation. Most models are based on calculation. You observe data, you calculate to a certain degree of thickness what a predictable outcome is, and then you sort of merge all these things together, telling you what an outsider perceives as what you're doing, which may or may not be right, and you make a model that will run by itself.

We have done this to a limited extent with some of our games. We had a game called "War in the East," which had an economic module for recreating the Russian war economy during World War II. It was a fairly large-scale game, and it interrelated with the cascading increase of Russian production because of internal factors — industrial organization or reorganization. But it also was affected by the overrunning and displacement of the Russian industry by the German Army, as well as by the loss and recapture of what we called Russian personnel centers, or population, because you could either run out of weapons or run out of men, or both, if you really blew it. This was an extremely boring thing to do by hand, and the game was so big that we wrote in RPG 2 — a 500-line program which recreated this complex situation. This was on a card system, too. Since then, we've gotten this big deal computer, but this game just sort of plunked along as we changed the parameters; we got the input when we expected them to lose Kharkov, Moscow, or whatever our supposition was; but the critical thing was that we were able to validate it through recreation.

To validate this system we had to be able to recreate what actually happened in the Russian economy as best we could determine. Now, the only sources of data, of course, were the historians and various other pieces of information that have surfaced since the war. So first we had to put together our model, so to speak, or our hypothesis of what actually happened in the Russian war economy, vis-à-vis what the output was, how many types of weapons, or how many actual units. When we had the system up we allowed the players to come out with different mixes of mechanized corps, tank corps, rifle divisions, rifle brigades, air units versus land units, and so forth and even increased their playing capacity. But what made it possible for us to get a fairly valid model was the idea of using history as a test. So, we tested it, and we do this even with our current games on current topics.

For example, the model we're going to use for the next war on the central front, we're going to start off assuming that it is the middle of 1945 and George Patton has suddenly decided to go after the Russians, or something like that. We have to start somewhere, with some premise. But, basically, the game has to work, because the geography hasn't changed that much, and we'll ignore the urban blight and what not for the moment. Actually, the part of Germany that

the attack was in hasn't changed at all; it hasn't been built up as much as the area by the Russian East German border. So, we will test the basic mechanics assuming it's World War II; then, one element at a time, will be fed in and validated and, of course, there may be as many as 50 or 60 separate design elements. From there on, it will get hairier and hairier, because we'll have no history to validate against except what we create in our minds as we go along to validate. But the point is that we will have a basic validation to go by.

In addition, we will go to what we call play testing in which we will have the customer, in effect, come in and test the game before we finish it up, do the final art version, and publish it.

Now the customer for your models is usually not the fellow who requested the product or who is the decision maker but some lieutenant colonel, GS-13, or someone like that down the line who will actually take it and perhaps decide whether or not anything's going to happen with it. Be that as it may, you can find a person typical of the type who will use it and have them come in an go through the prototype. You ask him, "Well, is this what you want?" and you tell him "These are the kind of results you may get." You can always be general because that way you can always fix up whatever either you did wrong or he perceives as having been done wrong.

We do the same thing with our customers, and this has been remarkably successful. We have 170 different games in print. They all still sell fairly well, mainly because we went through this testing procedure. First, they are modeled after history, and two-thirds of our games that are historical are quite simple. If you're going to simulate the battle of Agincourt, Look Out Mountain, whatever, you know what originally happened. It's very interesting when you do have to recreate in the game what happened, because the model can sometimes show you the way to recreate — the model can, in effect, become a part of the design team. It has requirements. If something is wrong, it can show you that it is wrong. In fact, I believe someone used the Battle of Chattanooga as an example of how this could not work in a model. Well, you can make it work. There you're getting into command and control. You're getting into what I call software.

I've often found myself at odds with the military and what-not, others on this aspect of software versus hardware, if you will. We have done a few small jobs for the Army and various associated organizations. The biggest problem we have is they're mesmerized by hardware. It's hardware over everything, when our experience has shown it's quite the opposite. Software is going to be the critical thing — software meaning tactics, doctrine, training, leadership, some of the things that were discussed this afternoon. This concept of software is the thing that's going to be critical. It's what will allow you to either call off Look out Mountain or the Golan Heights, or any number of other things. It is what allowed the Germans to smash France in two weeks in 1940, or perhaps it's what will allow the Russians to crush NATO in six days, if there's another war.

We cannot avoid these considerations. I mean, we cannot from a commercial point of view. Our customers are often knowledgeable in terms of the history or the current events that they're studying, and if we don't answer the questions that they ask they'll go somewhere else looking for answers. I think that you people are under the same kind of pressure, but because of the layers of users between you and the ultimate decision maker, it's perhaps not brought across quite as strongly. We also find that problem when we do particularly long-range problems. Most of our products take eight months from inception, from when we know we're going to do it to when it's actually published. When they take longer than that we tend to get a little lost, the games tend to get a little fuzzy as they go toward the end, so that may be a problem we have that you don't.

The final point is somewhat related to the first two, which were playability and testing. This point has to do with playing devil's advocate — having someone who will come in, either from your staff or somebody else's, and really rip you apart. We get this all the time. One of the hardest parts of our job is what we call rules writing. Now, this is something you fellows really don't have to deal with. In designing the game, we have to deal with algorithms — we don't use that many, obviously — and we have to deal with the game mechanics, the interfacing of all the numbers and the formulas and the data. But, the hardest part is writing a set of rules that somebody, or, actually anybody out there who buys the game, will be able to understand. We lose about half of our new staff we bring on every year, because they simply can't handle the rules writing business, and that's even after they've been warned that the attrition rate is 50% or better; that this is what it's going to be like and do you still want to try it. We have some very dedicated poeple, as you do also. There are a lot of people in this country who love to get into gaming, but we tell them this is the big problem. The devil's advocate in our business is not someone we have to go out and look for because we use our customers to test it, not just the game, but the rules. They'll come in and say "even I understand this." In fact, we'll even send copies of the game out after we finish this process to people who do what we call blind testing. We'll send it to selected gamers around the country who we know are real pains in the neck, very meticulous lawyers we call some of them, (some of them *are* lawyers) and all they will get is the rules and the game components, the prototypes and what not. They get no assistance from us, and the designer won't casually walk by while they're testing

and say "Oh, no, no, not that, it's this way." Then they send back a critique, and some of the critiques are heartbreaking. Five or six pages of tiny script come back in excrutiating detail. Some of it I can dismiss because it disagreed with my design decisions, but to a lot of it, I say, "Oh, my God, I left that out."

I'm sure you people have some problems like that, but since you are usually in attendance when these things are used, and since it's run by machine and most of the thing is automatic, you don't have as much of this problem. So you don't get the devil's advocate. But the devil's advocate will not just rip apart the rules as a means of communication but will often point out basic flaws in the logic of the game, or certain things that we assumed were working one way and not working the other.

A classic example of this was a game we did called "La Grande Armee." It was basically a strategic level game, a theater level game you might say, on Napoleon's campaign around 1805, in central Germany. This was about four or five years ago, and the fellow who designed the game had it tested. It looked very good; the rules looked complete; everything was fine. We got to the printing stage, and finally we got the tear sheets back from the printers. We sat down and played it and we discovered a very serious flaw. This was during the Auerstädt campaign when the French invaded and they defeated the Prussians near the border. The Prussians could always beat it right on down to Berlin and surrender. Well, the fellow who sat down in the office and just casually played it said, "Well, I don't have to do that. The rules don't say I have to retreat immediately. I can fight a delaying action. I have all these horses and all this calvary, I can just deploy them in a big screen and delay the dickens out of the French, which means I can rally the rest of Prussia behind me and stop them from defeating me entirely." Well, needless to say, we had a big errata sheet in that game, and it taught us a lesson: that you cannot let the designer of the game be the sole judge of how valid the decisions were in the game, or, more importantly, how valid the testing was in the game. This is something I think you folks are very much a victim of because of the nature of war, in which large groups of people are involved. There is the part of Murphy's Law that says when there's more than one person involved in a project no one can be blamed for anything. In our shop we have three or four people involved in any project at each level, actually ultimately maybe eight or nine, and I spend most of my time running around saying, "you are responsible." Sometimes it's somewhat capricious. Usually the designer says, "If this game goes down the tube it's your fault." I have a long list of all the games that have ever been designed, with their ratings. We rate them after they've been published although we don't go by sales much because they usually depend on the subject matter. A current topic game will do much better than a game on an older period.

We use more models in running the business than in designing the games — in regression analysis and correlation coefficients and so forth. I love it, but we can only get away with it running the business, which might be imperative. But, in terms of the game, we just ask people in bimonthly surveys, to rate the game on a scale of 1 to 10 for acceptability. This is basically their overall acceptance of the game as something that they wanted. We ask them "Was this what you wanted when you bought this game or when you originally requested this game?" The ones that go below six, and a number of them are mine, are just well, what you live with in business. We sit around in our weekly R&D and sometimes say, "Well, this one is a turkey." We have our criticism sessions; it's the only way we're going to find the flaws.

Now, of course, we have the advantage of doing about 30 or so new games a year. As I said, it takes us about 8 months to complete one, and we have about 20 or 25 in the works at any one time. We have about 25 people doing it, and it only takes us about 400 or 500 man-hours per model. Of course, our models build a lot on previous models so many of them are just variations of one model. But this gives us a chance not just to make mistakes but to learn from our mistakes and to immediately go on to something else. For example, the game I did 4 years ago called NATO I consider to be utterly obsolete. I was aghast when the Army War College called me and said they wanted to use it in their curriculum next spring. I went down there and tried to talk them out of it. Well, we compromised. I modified the game hastily, and they're going to use it. Hopefully it survives that process. Actually, it was quite suitable for their purposes. They mainly wanted to get the students involved in perceiving what the macro elements were in the theater; in other words, what all the divisions on the map looked like. That's an advantage we have. Our games use maps almost entirely, and they wanted the students to get involved in all the nitty gritty.

Now, some people here have mentioned the desirability of getting users to understand the basic elements in a model. If you could make, perhaps not a simplified version of a model, perhaps that's overstating it but one that requires some human interaction and with a little dickering could have it displayed on a TV tube, you could have the players start using a few of the elements. You just give them control of a few variables. You could give them a scenario like we do in our games. For example, for a given scenario, you could have a variable such as the attrition rate among ATGMs, or the attrition rate of Russian infantry versus NATO infantry, or any number of things, and just run it through a run. Now, since most runs can be done in 10 or 20 seconds on a third or fourth generation computer, depending on the model you're using, the fellow can sit there and punch in whatever changes he

likes, and zippity-do, there is the output. The output, of course, has to be fairly simple. You don't want a lot of mumbo-jumbo and formulas.

Even people who know what those formulas mean don't want to see them all the time. We have people coming in to play our games, who are scientists, physicists and so forth, and they can easily sit down and punch into a calculator. They're all knowledgeable about algorithms, but they find it's more fun to learn; or, they can learn more if it is fun. It's sort of a reward system. You can get them to use it more.

I suppose I'm talking about educational psychology: You can get them to use it more if you make it enjoyable. Now, some people will say you're prostituting it. Well, yes and no. If it gets people involved to the extent that they say, "I want more, I want more levels of detail, I want you to add this, I want you to add that," you can imagine what that can mean in terms of contracts. Somebody's got to do the work.

But if you're turning the people off, or, if you're not turning them on, let me put it that way, you're not going to get this interest in the models and what they can do. We've all heard the rumblings about the need for this kind of research work — do we really need it, does anybody really use it. These questions will probably be repeated as long as there is a military industrial complex. Well, the reason people don't use the games is because they can't use them, not because they're physically incapable or mentally incapable.

Half the population is mentally capable of using our games, yet, we only have about 100,000 customers. There is only about 250,000 people in the whole country who are into this type of game, as far as we can tell. But the potential is maybe 10 or 20 million, even if you discount segments of the population for a number of reasons.

For someone to want to use our games requires an intense interest in history, a bit of patience, and sort of a leaning towards a mathematical approach to things — models as it were. What we have to do constantly is to take out all of the mechanics in our games. We are now working on computerized versions, which will allow people to access our models as easily as you people. Even now with the hardware available you can allow your users to access your own models, and I think if you do that you'll be surprised at the results. You might learn more about them yourselves.

Dr. Hembold: What I would like to try to open up is an issue that certainly is of personal interest to me. It was mentioned yesterday by Jerry Bracken and was raised again by Dr. Anderson this morning — that is the question of whether the models, the gaming models, and so forth, do indeed predict the outcome, or if not, then really what is their significance, and how would the manager use the model results?

The purpose of the conference, as I understand it, is to focus on the use of these games in military force planning, which would involve the determination of force structures, force mixes, and force levels. It seems to me that the manager who is using this information must be in a position of predicting. He must be able to say, in effect, "if you field these forces victory will be yours," and "if you field those other forces then all is lost." It seems to me that in that sense the games, to be really useful to the manager, must in some sense predict the outcome in order to be able to help the manager to make that determination, and in order to help the manager be in a position of persuading other people to accept his determination and go along with decisions regarding force planning.

In fact, from what I can observe, managers do use the games essentially in that way. They do, in fact, interpret them as having a capability of predicting outcomes. Now, if they don't really predict the outcomes, I don't know quite where that leaves us. So it seems to me that we should either show that the games do have a predictive value, or else we sort of have to accept what seems to be the implication of the opposite, which is that the games are largely for purposes of self-education. The analyst himself gains insights, he uses them to screen some preliminary thoughts, he uses them to suggest new kinds of information that would be helpful or useful and so forth. But this is basically a use of the game to assist the analyst himself, who then may take some of these insights or findings and attempt to assist the manager or the decision maker.

Well, what would be involved in showing that some of these games do have predictive value? And in what sense is it possible to show that a game has some sort of predictive value? Well, there are various ways that that could be interpreted. One of the ways in which it could be interpreted is that the games, in some sense, correctly represent or correctly reflect the laws of war. By that phrase, I mean the physical, psychological, and social processes that underlie the dynamics of combat interactions. I was quite interested in Jim Dunnigan's comment to the effect that in his experience what he calls "software" seems to have such a great influence on the outcome of battles. I think Jim was thinking of "software" in terms of it being a factor that helped to distinguish one battle from the other. But I want to try to turn it around a little bit and argue that if indeed software is that important, then since people don't change very much, even when the weapons and the specific tactical items change, there may still be a large reservoir of what you might call invariance or persistent patterns that affect the outcomes of battles. As a minimum, that is one way of addressing the question of the validity of gaming. A particular gaming proposal or plan would be to see whether or not it does correctly reproduce those persistent patterns, those invarients. Of course, to do that you have to find them so that they can be used as a standard of comparison or as a test in assessing the

validity of the model.

Discovering invariants, I think, requires imaginative use of historical data. There has been some work in that area, but not very much. It has been largely exploratory and experimental; however, and as part of that work I think we have found some persistent patterns. These patterns tend to persist over centuries and generations, despite the advent of gunpowder and rifled weapons and other developments. That being the case, there's some hope that they are in some sense fundamental and that they will, because they are so fundamental, persist into the future.

Once these patterns are discovered or made known, of course, they should be continually verified against additional experience. It may be possible to do some work comparing field exercises and controlled experiments, in particular with new combat actions. I think that there is a possible way through this line of approach to at least, in some sense, partially validate gaming. I don't think this is necessarily the only line of approach, and I'd like to stimulate more thinking and more comments, and hopefully more work, more planning, and more results in this area.

Mr. Steenrod: My name is Stan Steenrod. I'm with the Army Concepts Analysis Agency. By way of a little background, I'm one of a few people at CAA who is doing theater-level gaming on something other than CEM. Actually, unless things so very wrong, we're going to be using IDAGAM 2 and a version of LULEJIAN in a study that is currently under way.

I tried to find my niche in this panel, possibly taking a little bit different direction from the other members here. Essentially, what I attempted to do was to listen intently to the speakers who spoke this morning and to write down anything that spurred my interest or imagination. Unfortunately, looking over my notes, I find they come out to a very disjointed set of comments. So what I think I'll do is select those comments that I have remaining on the last two speakers and just talk very briefly about a couple of points that I think are important.

The next to the last speaker was Mr. Asbed, my colleague at CAA, and he spoke about the use of the simple model there. Actually, I find I am encouraged, and would encourage others, to make more use of simple models. To a degree, you could even relate this use of simple models to some of the things Mr. Dunnigan just said about his models being simple and finding wide acceptance. Our experience has been the same. The study that resulted from determining the Army requirements for close air support, accepting the fact that we also used IDAGAM in the model, nevertheless had very wide acceptance and has become very popular. Whether you attribute that to the fact that we did what we said we would and came out with the number for the requirements of close air support, or whether you could attribute any popularity to the fact that there is at least in part a simple methodology that could be understood, you'd have to judge for yourself.

I might also give another example. I used to work for the Air Force, and in earlier discussions you heard a number of people take credit for the NATO shelter program in Europe. Now, I have no credit to claim in that area, but I do recall working with a colleague at Air Force Operations Analysis, who developed a very simple exponential model, which investigated this same problem. We programmed it in BASIC in something like ten or twelve instructions and put it on a little time-share terminal we had. The result was that there were a large number of general officers who were interested in the model, and the other individual was frequently requested to come around and brief his results and leave the model with the general officer. I think that was a clear example of something that an officer could easily do in a very few simple calculations to look at variations or sensitivities, if we want to use those words. So, as I say, I find the use and the appeal of the simple model very great.

I think you should also look, even in Mr. Asbed's presentation, and to a degree I can make this comment since I wasn't directly involved in that study, but you can notice the model creep, shall I say, even in that. We started with a simple formulation, which amounted to four unknowns. Before long we were talking about reserves, we were talking about redeploying the tank component of those reserves, and then we began talking about attack helicopter contribution, and about the artillery preparatory fire. We stopped at that point, or let's say they stopped at that point, but there has been some discussion since then of going to weighted weapons, all of which certainly sounds logical, also in talking about estimating FEBA movement. So, if this continues, before long this model will certainly be a competitor with the other big four.

The other point would be concerning some of the things Mr. Robinson brought up in his speech. I used to work for Robbie, so I'll tell him from the outset that I'm not at all being critical of the comments that he made. I have made one observation over the past several years with regard to adding command and control and communications, to our theater-level models. To a degree we've already circumvented that addition, you might say, in the data, but it may in fact be in the model; certainly a large number of our studies have always assumed perfect intelligence. In essence, what that amounts to is when you talk about having new intelligence systems, or new command and control systems to provide information to the commander, you find that you have to, in effect, backtrack and say, "will the FEBA really move farther in the game I ran last year because the intelligence wasn't really as good as I thought it was? And, now, when we add the intelligence, it comes back to where I told you it was last year." Not very many of your customers will accept this as being a very satisfactory answer to the addition

of command and control communications and information.

I think Air Force Studies & Analysis may have a better chance at success. They are putting their intelligence and command and control aspects decision making in an entirely new model so, consequently, they aren't faced with the residual memory of what the study showed last year and now the study is showing the same thing but all of a sudden it has more and better information in intelligence.

One last point is one that is relative to Mr. Robinson's comment on research. He talked about information, or input data from CPXs and also from manned simulations, both of which I support wholeheartedly. In fact, at one point, I not only tried but presented a paper on trying to couple a simulation model with the CPX so that one could run a CPX entirely closed, letting the computer play the opponent, provide the information to the various headquarter staffs, and so forth. This was suggested in a NATO arena. Unfortunately, at least from my viewpoint, there was never any interest shown in either the NATO headquarters, or the individual military headquarters where I tried to discuss this, but I still think some things could be done there. I think it would not only provide a more realistic CPX, but it would also allow us to capture data, which right now there is precious little way of capturing from the CPXs that are underway.

Professor Taylor: I'm Jim Taylor from the Naval Post-Graduate School in Monterey. I've been involved the past few years in teaching courses in an academic setting along the lines of methodologies and techniques of combat modeling. I'll address most of my comments, which I didn't hear addressed to any great degree earlier, to some observations regarding methodologies of theater-level combat modeling.

First, I think we have to start off by saying that theater-level combat models are embedded in the study process, and that there are different purposes for different studies. Depending on the purpose, different things receive emphasis either in the inputs to the black box, inside the black box, or what is generated as outputs from the black box. For example, one can look at system design, system trade-offs, concepts of employment, strategies and tactics, or operational planning. In all of these there are going to be different activities.

Also, maybe we should think one of the functions of the analyst, or intellectual community, as being the articulation of many of the issues. I think many times this is not done at all.

Why are we looking at theater-level models at all? Well, basically, combat is a combined arms operation, and if we're going to try to do any very dissimilar trade-offs between systems we're driven right away to the very highest level of operations. Also, there are very important resource allocations decisions made at the theater level that make it almost impossible to evaluate certain types of systems other than at the theater level. For example, one time I had a bright young Army officer developing a model of the sequencing of warheads for a firing site of an air defense system. After we got all the way through the study we decided that the battery was the wrong thing to look at. The reason for this was, with penetrator tactics as they are, the battery sees everything or nothing. In other words, the way to defeat air defense is to saturate and overload and break through, so, therefore, if you look at an individual site it's going to see everything or nothing. So, you really have to start looking up one layer above.

That brings up another point; why are we talking about hierarchies in combat modeling? Well, basically the military system is a hierarchy in which you start with the field army, and go down through many organizational levels. If you're modeling systems you're dealing with a very hierarchical combat system. It's only natural, and why is it taking so long to recognize that this is why we're driven to hierarchies?

Well, what about methodology issues? I think one great contribution that Jim Dunnigan has made is to articulate the point of playability versus credibility. These are things that war gamers have discussed for hundreds of years. When Livermore brought over the American kriegspiel, or invented the American kriegspiel, it was what was called a rigid war game with very very detailed assessment rules. It was unplayable. Throw it away, don't use it; then comes in free-form war gaming and this is something that is still around. There are lots of things that turn on military decision makers. Whether or not they're relevant is another issue, but there are many things that a decision maker likes to see incorporated in the model. He likes to see it enriched in various dimensions. Well, the more you enrich it, the less it is playable. I don't know what the answer is. Maybe the answer lies in an educational program. Jim's suggestion sounded like a lot of common sense. Give the user something he can do hands-on, so that he can understand and say "Hey, with a few variables I can do something, but with a lot of variables, or a very demanding data base, it's going to be less responsive."

So, okay, playability versus credibility. For the actual modeling part, something that I don't think was spelled out was just what are the significant variables. We can look at various things; military people will say what you do in combat is shoot, move, and communicate. Well, there are various levels of variables that reflect shooting, moving, and communicating. Okay, we've got behavioral variables, variables reflecting the behavior of either organizations or people. It doesn't necessarily have to be people; it could be organizations. When does a unit reach its break point? How are decisions made for committing of reserves, and various other things? Engagement of forces? There are time variables, of course, and space variables. I

learned something about our models in terms of spatial variables from playing Jim's game, and I'm glad to see that I'm not the only one who plays them solitare. Jim's rules are very complicated, and his game rules made me realize how complicated our combat models must be. One very enlightening experience I've had in conveying to students how complicated real world combat models are and why I can't bring a full-blown combat model in the classroom occurred when one of my ex-students here, who is now at SAGA, came up to me and said, "Gee, why don't you bring relevant things into the classroom?" You can't because they're too overwhelming, too much detail. You can brief students for a week on a model and not know whether you've really explained what the model does. To show how complicated our models are, let me illustrate with one of Jim's games, the Panzer Blitz. That is an excellent example for explaining to somebody how complicated DYNTACS is, for those of you who are involved in the land combat modeling of battalion level armor operation. You show somebody the rules for Panzer Blitz and it's an excellent pedigogical vehicle, but the rules are very complicated, I'll admit. I haven't really figured out all that you've got thrown in there Jim.

Mr. Dunnigan: Neither have we.

Professor Taylor: But, after you see that you say, well "gee, and that's the way the real world is, very complex." I think your games are excellent for portraying the spatial variations and the layout of the battlefield. I think that's always something that's hard to communicate. Probably in a study, there's a visual assistance or visual aid problem here. I've found in teaching courses, in quite a different sense, just showing somebody Lanchester equations, especially when for a fairly complicated system, I draw pictures of the various forces and draw arrows for who's firing at whom. That's very simplistic. In the same sense, I'm very intrigued by using the board game for laying out the order of the battle and for really modeling the spatial representation of these systems. I've just talked about behavioral variables, time, space, systems. We've got lots of different types of things on the battlefield. They're not just hardware, they're combinations of hardware and software. They're interlinked and structured, very hierarchical.

The next point is which of those things are we going to represent in the models and how? Models are representations after all. Then comes the question of how much detail, the level of detail — back to playability versus credibility.

Another point that I'd just like to bring up is understanding and differentiating between the conceptual basis of a model and the details of implementation. Here at this conference I've heard very little, at least in my perception, on the conceptual bases of the models. There are some significant issues here having to do with level of detail and aggregation. The use of index numbers, instead of very detailed models processing the outcomes of the engagements on the battlefield. Now, one thing that did strike me was seeing Captain Underwood's presentation on the naval campaign modeling is that that problem is quite different than the Army's problem in modeling ground interactions. In other words, the Navy's systems are very few but they are major systems. You model them on a one-on-one basis, two-on-one basis. The ground combat modelers, the guy modeling what the captain of land warfare is doing has lots of systems out there — hundreds of thousands of machines and people thrown over hundreds of miles.

Generation of engagements is another consideration; I think that the board game is an excellent vehicle for this. This last quarter I had students using Lanchester type equations for assessment, but using a map exercise for generating engagements in the Fulda Gap area. So I think one piece of the elephant in the combat model is the generation of engagements; here the board game is invaluable.

Then we go to the assessment of engagement outcomes. Well, here there are some significant methodological issues that have been around for a number of years, barely addressed, and they're going to be around after many of us aren't. These issues have to do with whether you are going to model on a very detailed basis, taking individual systems and tracking them, or classes of weapons systems, or whether you are going to aggregate through index numbers. It is not even clear that many of the major models ever tell analysts from the outside exactly what they do.

TAGS use differential equations combined with index numbers for playing the ground war. Although it has other options, IDAGAM uses, index numbers in the ground war. BALFRAM, which was talked about as a Lanchester model, really doesn't use Lanchester on numbers, but it uses Lanchester on firepower scores, and I think Seth (Bonder) brought up the point that firepower scores aren't worth a damn. However, the VECTOR models, which are very detailed, aren't used in studies right now, so we may be in transition. At one time, it was very critical on firepower scores, but then I used a pedigogical device with students, saying "Well, look, here's the menu of what you have available for techniques for modeling combat. You've got simulation, firepower scores, and Lanchester. I'll show you what each is, and that's it. The smartest people in our country haven't figured out any other way." Well, when I look at firepower scores that way and I put myself perhaps in Phil Louer's place, I'm not quite as critical. I realize that unless one can get very detailed and take ten years to develop a data base, I'm going to have to do some aggregation.

Okay, another point we might think about is, does one use one model or a suite of models? Does one use models in parallel, hierarchical, or in series? I think basically what the people from CAA have been telling

us is that it's very convenient to do modeling in parallel with a very detailed model to give testimony on the credibility of your simplified analytical model. I think we really haven't articulated the various options open to us.

Another point is to use a simple model to sell and illustrate points. And I think one thing that I want to bring up is that we really haven't talked about free form gaming and how it intersects with the very structured options given in many of these models. Or maybe I'm just reiterating something that Klaus Niemeyer said when he talked about putting the human in the loop.

Another thought that comes to mind is that when you talk about the behavioral variables, one way of modeling is to simulate. Really, what we're doing when we're putting decision makers in our interactive game is actually modeling the decision making process by having real people play roles in simulating.

I think one thing is that we have to try to understand the combat processes and the various phenomena involved. Things like break points, movement of contact zones, and how to represent contact zones. I think we should distinguish in our minds differences which I haven't seen here — differences between methodological problems versus implementation problems versus organizational problems. There are also organizational problems like institutional memory, and we have a continuing educational problem here.

I learned a good buzz word at the Naval Postgraduate school — Continuing education is in. From what Jim (Dunnigan) was explaining to us, I think we probably have a big continuing education business with the customer because we're playing musical chairs at the decision maker level in our business.

Mr. Dunnigan: Get them hooked.

Professor Taylor: Right. I've had the luxury of being able to sit back and watch from afar and not have my job on the line being involved in one of these things, or being able to reflect and become interested in what I'm doing, and looking at very idealized versions of this.

To wrap all this up, another very intriguing thing that Jim (Dunnigan) brought up is validation through re-creation. That's very intriguing. I don't think we're ever going to match up exactly. But getting a rough qualitative picture, this is a little more of a sophisticated approach, this is for the experienced analyst, but I think it can be done. I think many of the things that Herb Weiss brought up on things along this line in Lanchester modeling fall into this way of thinking and many of the things that Bob (Helmbold) has done with looking at historical patterns in battles, fall roughly, if not exactly, at least philosophically along that line.

I think the big problem is to develop a theory of conflict in combat. I say conflict because there are problems with modeling war initiation and also war termination.

War initiation, that may be a pretty weak link for us. You know, one writes scenarios, makes assumptions, but how much have we looked at historical data on war initiation? There are sociological aspects here. What I'm trying to say is that there are a lot of things to learn from history. The Japanese-Americans are very sensitive about Pearl Harbor, we know Pearl Harbor. Liddell Hart tells us that the Japanese sent their bright young officers to the English war schools during the 19th century, and one of the things they learned there was what a neat thing Admiral Nelson did when he caught the French fleet at anchor in Copenhagen and sunk it all. They went back and they got into a little contest with the Imperial Russians and they sank the Russian fleet at Vladivostok in 1905, and then they got into a period of rising tensions with the United States, and what do you think they decided to do? So, I think we need a theory of conflict. This has to do with studying behavior, also quantitative things; like what models for attrition, FEBA movements, etc.

I guess a lot of these comments were rather random, but I've tried to tie some of these ideas together. Obviously, I haven't answered any questions, but hopefully I've given some points for us to reflect on. I think the last thing to be aware of is that these are not simple problems. We're talking about very complex things. The only thing that's really definitive and clear cut in this whole business is the answers of our models, not the real world. Colonel DuPuy tells us that in most of the World War II battles if you look closely it's hard to tell who really won, but yet our models give nice clean cut answers.

So, with that, I'll finish up. Thank you.

Mr. Farrell: Thank you. We have time, I think, for a few questions for the panel or for previous speakers.

Questions: I have a question for Mr. Dunnigan. I was interested in your story about the Prussian delaying action and one of your employers refusing to retreat and surrender requiring the issuance of an errata sheet to go with the game, which leads me to ask what is the purpose of your game. Isn't this a legitimate tactic that might cause history to come out differently?

Mr. Dunnigan: Yes, you're quite right. The designer made a very grave error in playing it out, trying so hard to make it recreate history that he did not allow for one of the more obvious elements that could change. On the other hand, in some games you could disallow the ability to retreat because one could argue, depending on the situation. It's quite possible that there was a very low probability of the Prussian Army recovering from the shock of the defeat they received. However, it was our determination.

Question: There was no possibility. The Prussian Army was crushed.

Mr. Dunnigan: Well, yes, but in the spirit of our

games we allow people to deal with any number of variables. In the strategic game you don't necessarily get defeated to the extent that it happened historically, and what happens in many of the games — take the Prussian player for example — was that this player would not commit his forces to a duel battle in the idiotic way the Prussians originally did. Indeed, they would get beaten on the frontier, but perhaps with only a portion of their army, leaving another portion still intact. The intact portion then, could make a fighting retreat back to Berlin, so for that reason we left it in.

Question: This leads me to my other question. That is, what is your experience regarding players replaying the same game?

Mr. Dunnigan: On the average, they play the game about 4.5 times. We do a lot of surveys, so I've got a lot of that down cold. They play the first time, the historical scenario, and they might play that once or twice, and then they'll try it a few more times using some of the variations and that's it. They'll spend about 20 or 30 hours on each game, which is more than you spend with a book, so I'm very flattered by that.

Dr. Huber: I'd like to comment or question the proposition that we ought to include more behavioral aspects, or let's say doctrine, into our models so that our models really reflect what current military would do. This has special pertinence if we use the models to, let's say, deal with events that are in future, maybe ten years, maybe fifteen years into the future, because we'd like to get some answers to what we are going to do in terms of our own objectives, in terms of weapon developments, etc. Because it comes to me that if you really look at history, you'll find so many incidents where the military weren't really capable of exploiting the technology at hand. This, I think, shows you how important it is to have the doctrinal constraints on the optimal use or efficiency of the weapons.

Mr. Dunnigan: Yes. This, I think, shows how important it is to have the doctrinal constraints on the optimal use of efficiency of the weapons.

Dr. Huber: Right, that's what I mean, you shouldn't include too many things that are present doctrine. You ought to use your model to investigate doctrine in terms of trying to optimize.

Mr. Dunnigan: Well, exactly, but first you have to show the fellow what the weapons can do the way they are constrained now, then you give him a model and say "All right, this is the way we do it now." Say you change this, blink, blink, blink, you say, "Look, you get a 50% higher P_K, and the reasons for it are this." If you demonstrate it, then they'll believe it. People believe what they can see and the more obvious the better.

Professor Taylor: On the other hand there may very well be some behavioral variables that don't change that much from war to war. There was an interesting article in *Fortune* in 1964 written by an Air Force colonel against computerized war gaming saying, "my God, look, people fight battles, not computers." It's an interesting thing for us to read today. He was talking about dangers of computerized gaming which we have carried a whole order of magnitude further since he wrote his article, but he brought up some very stimulating points that still make interesting reading.

Question: I was fascinated by the reference made by several of the members of the panel to the fact that this type of work that you do reveals what's described as the laws of war. I'd be interested in having some analysts describe to me what you have in mind when you talk about the law of war, and also whether or not there's any consensus within your community that all of the work that you have done, say over the last seven years, has revealed certain new laws of war which are unique to the particular period.

Professor Taylor: Me?

Question: Yes sir.

Mr. Dunnigan: Well, I didn't use that term, so I'll just give you my version. The principles of war I call them. Economy of force, secrecy, surprise, concentration. I don't know of any laws of war. The only law is to beat the hell out of the other guy, seriously. I know what principles are, is that what you mean? Are they synonymous?

Question: I don't know, I was just asking the panel.

Mr. Dunnigan: Well, is that what you mean? You mean the principles of war?

Question: The first speaker spoke of laws of war.

Dr. Niemeyer: I really didn't say laws of war. I really said insights into the phenomena of the war. But if you want to have a very good law of war it's the Lanchester equation, for example. This describes very well one type of an assessment procedure related to certain specific situations. I think this is one that is pretty well known.

Colonel DuPuy: A demonstration of the principles of massing and economy of forces.

Dr. Niemeyer: Right.

Question: Do you have a feeling that in all the studies that you've done over, say, the last seven years, since the turn of this decade, has revealed, or is there a consensus, that certain things that you've discovered were not known previously?

Mr. Dunnigan: No, I don't think so. I'm a firm believer in there's nothing new under the sun. I've said this before, we steal most of our ideas.

Professor Taylor: So do we — from you.

Mr. Dunnigan: Yes, exactly. I'm a great admirer of Johann Bach. He was not a great innovator, but he was a great plagiarist. He just did a lot better with what existed than anybody else had, and I feel if anyone can do that they're doing quite well. Our main contribution is to make available to more people more examples of what has gone on before, these basic laws of

the universe, so to speak, like Lanchester equations, or what have you.

Question: So, really what the studies seem to get at then is to see how the traditional laws of war would apply given a certain set of circumstances?

Mr. Dunnigan: Yes, the problem is one of education I think. Also, while you're doing that you're constantly questioning these laws of war. I mean, they're not carved in stone, they're more like made of jelly and they have a tendency to change shapes slightly from time to time, and if you don't keep up with the changes in shape you're likely to get caught, you know, in the wrong posture, shall we say. This is one thing that I think this constant study does, and I think this is the one problem you folks have. You do the most powerful study — I've seen some of the work that goes into it — but a lot of it becomes obscured. It's like the tree falling in the wilderness, nobody hears it.

Question: Well, what I was going to comment on, and what I thought somebody might mention, is that I recall a good deal of literature in the not-too-distant past that was trying to make the case that the blitzkrieg was dead.

Mr. Dunnigan: Not us.

Question: That formal warfare, because of technical innovations in weaponry and other factors, was no longer a practical way to . . .

Colonel DuPuy: The Blitzkrieg isn't dead, it's . . .

Mr. Dunnigan: Right. I'm a great fan of VonHutier, the fellow who invented infiltration tactics. There was a hopeless situation for the offense, and he found the solution, and I think the same applies today. It was in the popular press, this nonsense about the blitzkrieg being dead. The blitzkrieg is there as long as somebody's willing to look hard enough and pick up the slight changes in the jelly that pass for the iron laws of war.

Professor Taylor: That's very important when we get into the resource allocation problem. Where I've looked at a lot of time-sequential, highly idealized allocation processes you find out that they're very sensitive to variations in how you quantify military objectives, how you model various things in the combat process, and how you even model things like the combat termination process. Once you get in a little farther, you find out that if you really optimize in a mathematical sense, the thing becomes so complicated that it transcends human comprehension. So there's another dimension — the very simplified board game that portrays the information that a corps commander actually sees and uses, that reflects his intuitive feeling for things is, I think, a learning mode. What role will optimization play? What role will just simulating human decision makers play, and what are their roles across the broad spectrum of the analysis and the study of conflict? These are some of the questions we should think about.

Prof. Taylor: No, I think many of my students . . .

Dr. Helmbold: Excuse me, but several speakers, and I was listening very closely for exactly that point, alluded to exactly that. They said well, we're going to do this so we can find out what happens, and several speakers said that or words to that effect.

Dr. Anderson: We have a problem right now with someone thinking they're predictable and they're not. Who?

Col. DuPuy: The majority of the people. We're caught up in our own war games.

Prof. Taylor: Yes, exactly.

Col. DuPuy: We believe them ourselves.

Dr. Farrell: We have time I think for one more question.

Question: Dr. Helmbold you indicated that the games should be predictive, and, as I understood it, you offered simple ways in which the community could move in that direction. At the same time you indicated that the users already consider them to *be* predictive. Now, the question I ask is what do you consider to be the cause of this misapprehension and what's being done by the modeling community rather than attempting to meet this expectation at some unknown date in the future, to prevent and negate the dangers of this misapprehension?

Dr. Helmbold: Well, I'll take the last part first. I don't really see that anything is being done to correct what I consider to be a misapprehension. As far as the results of the model being considered as giving predictions, that was the reason the model was done in the first place. So if it doesn't in some sense predict an outcome, the purpose of the whole exercise kind of goes a glimmering and leaves the exercise really kind of empty. That they are considered in this way really means that at least some of the time they're considered to be credible, but just because something is credible doesn't necessarily mean that it is valid.

Dr. Anderson: Or perhaps, nobody really believes they're predictable. Maybe we should name a decision maker who is under this misapprehension that you're talking about. One person who thinks they're predictable. Do you think they're predictable right now? You said not, does anyone else? Well, maybe that really isn't the problem.

Chairman's note: At this juncture in the meeting, amid considerable hub-bub and contention, the Session Chairman called for order and a coffee break, noting that we were in severe scheduling difficulties with respect to the start of Session IV. I note this point because Bob Helmbold had identified an issue over which there was obviously significant disagreement as substantiated by the Session III closing remarks of Bruce Anderson, Jim Taylor, and Trevor DuPuy. I was subsequently asked during the coffee break by Bob Schneider (Chairman, Session V) and others to delay the start of Session IV so that the debate might continue but this was not possible because Martin Shubik, next scheduled as the lead-off speaker for Session IV, had prior commitments and was on a very tight schedule.

The issue raised by Bob Helmbold is, of course, an explosive one because it is precisely over this matter of believing in the predictive qualities of models that the analyst stands accused by many on the fringes of the profession. Here, however, was the case of an analyst, and a highly respected one, suggesting perhaps, that his colleagues stood justly accused in overselling the products of their labor. I had subsequent informal discussions with Bob Helmbold and others involved in the final flurry of Session III, both during and after the meeting, and have concluded that what appears on the tapes concerning this issue is beset with an unfortunate combination of semantic difficulties, obtuse expression, and overstatement. However, the fact that it did surface at all as an issue I believe to be significant. Further discussion of the matter appears in Vol. II, "Summary Discussion of issues and Requirements for Research".

Due to the press of time, Professor Shubik continued as the lead-off speaker for Session IV without the benefit of opening remarks or an introduction.

Session IV Game Theory in Theater-Level Modeling: Optimal Solutions and Heuristic Solutions

22 — The Role of Game Theory in Force Planning and Game Associated Problems

PROFESSOR MARTIN SHUBIK
Yale University

Overview of game theory literature and research and some sage observations on its application potential

Prof. Shubik: We're really not up on all of the current theater models. I also find that there's a very dangerous occupation with abstract models which is X and original sin, X and the Polish question, X and something else. You take an abstract model and you give it a new name, so yesterday's tactical model is today's theater model, and you can't tell it without reading your program. As such, I am extremely suspicious of hanging titles on a specific abstract mathematical model unless along with the titles you hang some very explicit hard modeling explaining why you claim this model is at one level of aggregation or that model is at another level of aggregation or disaggregation.

Now, let me get straight to the point. I am a confirmed addict and believer in the value of game theory, but I'm at the same time a confirmed believer in the dangers of misuse as well as use of all types of models. Explicitly, there are three terribly important values to be attached to modeling at as formal a level as game theory. The first is as an antidote for spurious reality, the second is as a logic check, and the third and possibly most important is a sorting device for relevant variables. I've said this particular point on many occasions and I think I will say it again. One should encourage one's staff to butcher lots and lots of cheap simple models. One of the ways in which you encourage them to do so is to make sure that the lowest grade of yellow paper pads are supplied to all of them; Royal Selectric typewriters are kept far away. If they belong anywhere they belong in sales organizations, and for those of you who are both modelers and have sales organizations, keep your Selectric typewriters in the right place because as soon as you use quality bond paper and fancy type face you start to believe unbelievably bad models. So the general point is model fast, model simply, and throw a lot of them away. But the reason for doing it is that they are logically organizing conceptual devices to enable you to start to sort out variables, and variables that apply in one area may be totally irrelevant to other areas. Among the other standard words of wisdom, but one that I believe to be so absolutely true that it's worth saying again, is that an all purpose model is a no purpose model. Therefore, the question and the ability to model pertaining to the question is very much the name of the game.

Now, I want to give you first of all a very brief, and I hope more incomplete than I believe, sketch of the theory of games as applied to military operations research. There are really a three by two set of distinctions that you want to make. The first three are two-person zero sum theory, two-person non-zero sum theory, and n-person theory. The other part that gives you the three by two is static and dynamic. Now, given those, I am a firm believer that one shouldn't even think of game theory models without keeping in the back of one's mind gaming simulation and other analytical models. They have to always be looked at in terms of the interaction in that direction, and I must also admit that I was extremely glad to hear in the previous comments some of the observations made on playable games. I just got back from Germany a few days ago where I gave an address to a group of economists in which they looked extremely puzzled because I offered to them two, not one, rules of modeling. One rule was the game theory modeling rule which is logic, consistency, and completeness. The other rule was playability. The reason I stress this time and time again is that in the things that we're interested in we expect that the system we're studying is going to be operated by human beings of a class not quite equivalent to John VonNeumann. It would be nice if we had a large supply of John VonNeumanns, but we don't have a large supply of John VonNeumanns and the playability criterion is extremely important.

Now, in two-person zero sum games, sort of the trinity of reasonable uses has been allocation problems, duels, and games of pursuit and search. For allocation problems, a good place to look, as many of you undoubtedly have, would be Dresher. There is a nice literature on Colonel Blotto games and some of you may be aware of this and in a little while when I talk more about the future than the past I want to suggest that I think there are a lot more twists that can be put on Colonel Blotto games which could be of some use. When you look at the literature you've got that class of games, you've got Blackett, Dresher, Beale, Hesseldon, and several others. Then, at the two-person level there was the work on tactical air support. You've got Caywood, Thomas on fighter/bomber, Berkowitz, Dresher on air support for ground forces, Fulkerson and Johnson on tactical air war.

Now, I think that all of this work was pretty high class, but somehow or other, and this really involves more the sociology of science than science, it more or less seems to have withered on the vine. It had this great growth in the late 1950s, early 1960s, and then every now and then you pick up an extra article, but it just stopped. The strange thing to me, at least, is that the people were good and the work, in my opinion, started to show some promise, but it managed to get shunted off and just didn't continue. I hope that I'll have time to come back to that.

Well, the other aspect where I think game theory has faired possibly better has been duels. Now, there are really several questions, because once you mention duels, why go the game theory route anyhow? Why not go Lanchester equations? Why not go Richardson's models for that matter? Dynamic programming models? To me it is still a very very open question as to which approach is better or worse. I frankly think that one of the reasons why it's an open question is because there's no single answer, and it just depends explicitly on what aspect of conflict you are trying to study.

The difference basically is this, that in some sense one of the reasons why I like game theoretic models is that at the very least they serve as a good negative antidote to sloppy modeling. In other words, in general, I don't believe most game theoretic models of conflict, but I like to drink my rational behavior potion pure. In other words, what I'm saying is that one of the good things about, say, a two-person duelling model is that it doesn't try to fake spurious human parameters that haven't been measured. Therefore, it doesn't define indefinable degrees of cooperation, reaction parameters, etc. which you find in the boys who push around differential equations, difference equations, dynamic programming models. Richardson's two books, I think, should be read in this

respect. Rapoport has a little book called *Fights, Games and Debates*, which I find is excellent as an antidote to difficulties and problems in modeling human combat.

Well, I think that in the area of jewels there is, as many of you know, the strictly mathematical work of people like Bellman and Gershik, Belzer, Karlin, Shapley, Smith, Sweet. They do, I think, on the whole, pretty high class work. There was very little attempt in that work to interpret precisely whether you could regard it as one-on-one or whether you wanted to regard the units as aggregated. In most of the mathematical formulations, there is just something called the player or one side, and you have to do your own hand waving to decide what that particular player or one side is.

In the search and pursuit games there is some antisubmarine warfare stuff. There are now several books on differential games. You've got Friedman, Blaquiere, Rufus Isaacs, and last but by no means least, an enormous Russian literature. As a matter of fact, if any of you are at all interested, a shock that both Bill Lucas and I shared more or less simultaneously was when Professor Vorobyev presented us with the Russian bibliography on game theory. I mean, it included non-Russian as well as Russian items, but it ran if I remember correctly I think it was 2,100 items. Yes, 2,100 items of the really first class bibliography and the Russians have been doing a fair amount of work in this area.

Okay, now, in two-person non-zero sum game theory there have been at least four areas of work which I find have been of some interest. One was the inspection problem which was worked on mainly by Maschler. Then there is a lot of what you might call soft literature which I feel should not be totally ignored. It was like a fashion show. At one time it was so overrated that it was the chic thing. It was sort of the okay stuff in the days of Camelot. Now it's sort of been wiped out and what I'm referring to is sort of the soft literature on escalation and threat and that includes Kahn, Schelling, Wohlstetter, myself, and a few others. I think that it was oversold at the time, and now it's being vastly undersold, because at least the people were worrying about the characterization of some of the major aspects of the preliminaries to warfare and warfare itself at theater level and above.

Another little art form which some of you are probably acquainted with was Norm Dalkey's work on nuclear exchange, and that particular model was a two-person non-zero sum game and was treated in a fairly interesting manner.

Now, as far as I can see, and this may just be confession of not having done as careful a literature search as I like to do, in the last three or four or five years the amount of decent military game theory published has been extremely small. You can check the output of MORS, the *Naval Logistics Research Quarterly*, ORSA and TIMS, and scratch around elsewhere, but at least to my knowledge, I've really been surprised how little is to be found.

Yet, at the same time I would like to note that the actual store of knowledge in fairly decent game theoretic results has been shooting up like crazy. The nonmilitary aspects of just straight game theory, and for that matter even experimental work, has been mushrooming. In particular, what we seem to have at the moment is first of all a relative isolation of sort of an almost ossified art form of the two-person zero sum game as applied to duels and occasionally implicitly or explicitly built into weapons evaluation models, virtually very little two-person non-zero sum theory being done in a military sense, and essentially no n-person theory.

The other thing that seems to me to have happened in the last few years, in general that is, I'm not really talking so much about game theory, has been the isolation of different professionals from each other. Now, I suffer from having been a consultant in the halcyon days of Rand, but for better or for worse it was genuinely real. The people in the math department and the economics department and the social sciences did talk to each other. The flow of U.S. Air Force colonels through for the logistics gaming lab was a real phenomenon, and they talked to each other, and there was a serious cross discussion. At present, at least from what I can see in the military operations end of things, the great achievement of the last six or seven years has been to separate out gaming from simulation and to almost kill gaming in the process — the great rush to the all-computerized model. And, at the same time, game theory itself has just been isolated and sort of left in the curiosity box.

Now, another thing which I call to the attention of all of you in this room, is that the age structure of the idustry, to my mind, has become extremely unhealthy. When one draws a graph of the age structure and you look at the number of individuals between say 26 and 34 or 35, they're getting to be pretty few and far between, and when the industry was healthy it was the other way around, so we really need to take a look at that.

Tied in with that, it comes as no news to anyone, I think, although I hope somebody can correct me, that the depth of interest in the war colleges in game theory and in serious gaming is pretty low. Now, when I say serious gaming, first of all I should cut gaming up into at least a couple of different portions. There are the sort of gaming exercises that have become almost like raising the flag in the morning, so you're going to do them come hell or high water. They're part of the curriculum, and, by the way, in the explanation of any organization a good analogy to remember is sort of the problem faced by a television company or a radio station. They're got to fill X hours a day and the question is how much of this X hours can be filled with ease. I always claimed that much of the teaching process anyhow is show biz, and if there's anything you like it's a good canned program that's sort of tried out and is not going to rock the boat too much. So once you get a certain game down it becomes like a Japanese tea

ceremony. Everybody likes it. It's got just about the right amount of honorifics, and you can go through the same hoops for a long long time because it eats up hours in a respectable way.

Now, I think that the state of gaming is, in fact, worse than the state of game theory. Because, I do know that the body of literature of game theory, although not in military applications, has literally been exploding and there is every indication that it's going to keep doing so, and sooner or later we're going to tack in on it. To me, it is extremely comforting to know that there is some area where intellectual repletion and the building up of reserves is in fact taking place. So in spite of the fact that we may not be using it at the moment, it is there. But, I feel that you cannot say the same with many of the aspects of gaming that we might be interested in for the simple reason that I believe that the fashion show for example had the political military exercises okayed, then the fashion show more or less wiped it out. I mean, it still survives, but on a scale. Already discussed here has been the switch from manned machine to machine and the whole question of the tendency to go inside the computer. Actually, I am hopeful and I sort of pick it up even from this meeting, that it is possible that the tide is beginning to turn. I think there are just now enough people sufficiently worried about the black box that maybe we'll start to swing back somewhat in the other direction.

I would like to note that in nonmilitary circles there has been what I regard as an extremely healthy buildup in the last few years in gaming and experimental game theory. In other words, there's now a fairly large literature of experimental games trying to validate game theory solutions, and, with due lack of modesty, at least two small bibliographies on this appear in my books in *Games for Society, Business and War* and *The Methods and Uses of Gaming*. I'm happy to say that although they're only two years old they're already out of date because I could double the size of the bibliographies in that area. There's been some not so good work, but highly relevant to us from a different point of view, which is experimental gaming and game theory in social psychology and to some extent even political science.

There is idea of trying to build, for want of a better description, what I will call Mickey Mouse games to experiment with just to check out very crude concepts of threats. But the social psychologists get gung-ho about personal differences. As a matter of fact, the fundamental difference between the social psychologist's point of view and the game theorist's point of view is that the variables that the game theorist and the gamers frequently sort out of their models are precisely the ones that the social psychologists concentrate on, such as the difference in the quality of personnel, sex differences, racial differences, etc. For example, take the prisoner's dilemma game played with male subjects versus female subjects, and there's a lovely little literature showing with dubious validity that females tend to double cross far more than males. I just note it to you. Maybe it's even sort of worth looking at.

Okay, I could ramble on, but I've just tried to give you the briefest of summaries of the state of the art. I'd like to switch to first of all my beliefs, then talk a little bit about tomorrow, and close with some specific areas that I believe would be worthwhile to work in.

My first belief has already been expressed by others here, and that is to stress simulation at the expense of gaming and game theory is highly dangerous. The next belief that I have is that there is a problem of the sociology of the trade, although in the last five or six years maybe it has become just a problem of age and the age contour of the profession, but there is a real need for better communication between the military, the theorists, the model builders, and the model operators. I think we've really got to look at ways to improve that.

The other question — it's the perennial one that comes up — so I feel I probably said this ten years ago, but I'll say it again today: everybody gets up and bows to validation and sensitivity analysis. I always like to add, what in the hell is the meaning of validation? Validation for what? Sensitivity analysis for what? I think that these are still open running sores and they have no simple cure. But it is terribly important that before you go through a whole elegant set of statistical tests, you should remember what you're validating for.

I cannot help it James — many years ago Jim Mayberry, at a gathering such as this, said, "Validation is a satisfied customer." And, you know, that wasn't such a bad definition. As long as you're not going to be the test pilot in the plane that's got to be flown, as long as they want to accept your blueprints that's fine. In fact, there used to be a great old cartoon which I wish I had with me which showed you a gentleman looking up in the air and there was a plane about to crash and he looks at his colleagues and he says, "oh well, back to the drawing board." So, it depends on what view.

Now, I believe that there is need for more experimental gaming combined with game theory and I really believe that it's in the military self-interest to start thinking about what small areas would be worthwhile to start to work in. The whole question of the concept of threat is particularly a good one. I also am of the opinion that the stopping of application of analytical game theory in military affairs, or sort of the attenuation, was premature. As a matter of fact, at this moment, because there has been a large buildup, I think that one could spend some pennies quite valuably in that area.

Now, I want to close with a few specifics. The first thing that I want to note, and I could argue that this particular example really, at one level, could be interpreted at least at the level of the theater, is a new twist to the Colonel Blotto games, which I call "extended Colonel Blotto games," or "Colonel Blotto and the insurgents," or

"Colonel Blotto and the saboteurs" if you wish. I'll assume that at least half of you know the Colonel Blotto game, and I'll make these quick remarks about it.

The general idea is this, the Colonel Blotto games had a certain number of isolated targets and they were attacked, and, depending on whether one target fell or one target didn't fall, you scored something. Then you added them all together and you get an overall score for the attack and the defense. Now, it seems to me that one of the absolutely critical areas for our work happens to be in the defense of integrated systems, and this includes whether we're talking nuclear warfare, whether we're talking sabotage in large-scale systems, whether we're talking power failures in networks, like command and control problems. I can name a lot of them. Well, here you consider a network with n nodes. Now, for those of you who know a little cooperative game theory there is a function called the characteristic function. The characteristic function has the property that if you consider some subset, S, it gives you a value for S, some particular number, and if you've got n nodes you've got 2^n subsets that you can fool around with. Okay, what you can use the characteristic function to represent is to not look at the n-person game as an n-person game but to look at the characteristic function as a description of the complimentarity in a network. In particular, for example, you could have the whole question of redundancy in networks come in where you might, at the simplest level, have a characteristic function consisting of just zeros and ones which simply say that as long as a certain number of nodes hasn't been captured that system still functions. But as soon as you go over a critical number of nodes that system is knocked out. This immediately gives you the possibility of doing an analysis in which you start to consider rather delicate complimentarity in systems. And you also consider in a direct way the trade off between the need for defense forces and the building of redundancy into a system, because one has always got the choice of whether it is cheaper to overbuild the network so you can have six or seven nodes knocked out and still have the system function, or to build it at a much smaller level and defend it like hell.

Now, a quickie simple example of such a model would be to let there be force one, one side has A units of troops, side two has B units of troops. Then there are several ways of doing the various versions of the Colonel Blotto attack on the different nodes. A quick and simple version, which is what I would call a continuous Blotto and the insurgents would be that he splits his troops to defend the n nodes, and the probability that the i^{th} node will fall is $\frac{A_i}{A_i+B_i}$, which is the ratio of the troops assigned to attack different nodes to those assigned to both attack and defend the boxes.

Now, when you have the A and the B very nonsymmetric, so one is much bigger than the other, you now can start to use this type of analysis to try to figure out the first requirements required to get yourself a probability of survival beyond any particular level you want, against a certain size of attack on a network.

Well, I just wanted to note this as a possibility, and to note that you can ring the changes on the characteristic function to sort of reflect any level of redundancy in the system that you want to consider.

Now, another area of work which I think would be highly worthwhile and technically not that difficult is to do a great deal of the duelling work again. But do straight game theoretic solutions on duelling work with the addition of a breaking point parameter put into the models. It can be done, it really isn't that hard. It just means that the strategy space would change. It's as though you're downgrading the worth of the different sides. I say this because at least in the couple of occasions that I looked at empirical data, the breaking point phenomenon really seemed to be a historically sound phenomenon. There's a saying, by the way, from the English House of Commons which is, "the arguments of my dear colleague are sound, mere sound." So, I'll leave the argument as to how sound the breaking point hypothesis is to those who have more experience than I do, but the point that I wanted to make is that it is not that hard to modify fairly straightforward analytical models to pick out phenomena like that, and if you do it parsimoniously it's sort of worth the effort.

Now, the next thing which I have seen done in very very few models is sensitivity analysis with respect to information conditions. Explicitly, game theoretic analysis shows how horrendously sensitive many of these models are to those conditions. A little example which I did (I never published it because I never finished it and I just lost the energy) was to take the Dresher/Berkowitz model of tactical air support and just change the information conditions slightly. It became a totally different model. If you really want to think about how we go through life, sort of saying, "oh, yeah, information conditions are important," but then having bowed in that direction we forget about it, I would like to remind you to think about what subscription any one of you would pay for one day early delivery of the Wall Street Journal. That does seem to point out the sensitivity of some systems to minor changes in information conditions.

I really feel that it is not that impossible to start to attack the sensitivity analysis on information. This leads me to my next problem, which is nonsymmetric gaming. I think that we have a dangerous tendency to concentrate too much on symmetric models. This isn't completely true. I have seen some nonsymmetric models. But I really feel that it is extremely important, and especially in modeling, say, US/Soviet conflict as an example, to take a parameterization, even though it is a terribly crude one, that picks up, at least at a first cut, the difference in bureaucratic parameters, political parameters, and economic parameters. "The Guns of August" on the Soviet mobilization are a lovely example of the slowness and the difference of bureaucratic parameters.

This leads me to my next point, which is the interface between military gaming and bureaucratic gaming. I think that that is an absolute key element. Some of you are only interested in the four-day war and the nuclear exchange, and I've got no argument with that, but there is a lot of military gaming which I think depends in an extremely important way on the bureaucratic interfacing and the flexibility of the decision systems.

The next point is, and I don't like the word because it's a pop word, but I use it and that is "fuzzy" gaming. What I'm saying about fuzzy gaming, and this is again essentially related to sensitivity analysis, is that in trying to design games with malice aforethought in which there are sort of bland probabilities connected with the lag in information, in other words, fairly flat probabilities of failure to activate even though an order or a command has gone out. Here, there are at least some probabilities of misinformation. I really think that the question of foul-up gaming is something we haven't spent enough time with. I was terribly impressed at the Naval War College awhile ago to see a great dog and pony show related to the Marines. It was a Marine film, showing their landing. It was amphibious warfare, the decibel level was above the Concorde, yet everybody was seated at impeccable desks in their command quarters, and the command and control was just impeccable — the messages were going out, there may have been three pieces of paper in every in-box and the out-boxes were clean. It was just marvelous.

What I really do hope is that we think seriously, even at the cost of a million bucks or so, of getting back, with caution and without too much overemphasis on soft gaming, to political military exercise gaming, and to pull back to the picture, a few historians, a few political scientists. Even if some may regard it as a waste of money, it will keep us a little more honest and the game a little more open, and it starts to stir another look at what parameters might count.

My last comment is that I hope that, if we have another meeting five years from now, the age level of our successor group at that meeting will be 5 to 10 years younger than the average age level of this meeting.

Prof. Lucas: If anyone wants to try to ask Professor Shubik a question, go ahead. *(exit Professor Shubik)* Jim?

Prof. Mayberry: I'd just like to add that Martin (Shubik) omitted a word from my comment of what a valid study is. It wasn't only the one that makes the customer satisfied. It's one that makes the customer satisfied after he has become informed. If it makes the informed customer satisfied, this is essentially the difference between swindling the man and trying to persuade him to be a good customer.

Prof. Lucas: Well, we've already lost our speaker I see, he sort of gave his talk and ran.

Opening Remarks (Displaced)

PROFESSOR WILLIAM LUCAS
Cornell University

I would like to stress the point that Martin (Shubik) made regarding the fact that there has been a great deal of theory in game theory in recent years. I think the expectations in game theory have oscillated drastically in the last 30 years. In the late 1940s, there were these great expectations. In fact, when the Von-Neumann/Morgenstern volume appeared during World War II, there was apparently serious consideration given to classifying it. We didn't want it to get into the hands of the Germans or the Japanese at that time. That's to be considered perhaps a joke now, but the expectations were skyrocketing then. But by the early 1950s, I believe they kind of plummeted. People found out games weren't all zero sum, there were some mathematical difficulties and various other problems arose at that time. But I do feel that from the late 1950s, that is, 20 years ago, there has been a lot of research. Several very rich subjects have developed and those who are informed, I think, have had rising expectations for the subject but with a much more modest slope on the curve than had happened earlier.

Obviously, the organizers added this session to the meeting, so they must have some hope for this area. I would like to take just one moment to mention one person whose faith in the usefulness of the game theory never did wane and that was Oscar Morgenstern. Oscar was one of the cofounders of game theory; he died just two months ago. Many people here knew him and he'll be certainly a loss to this community. He was kind of an untiring consultant and consistent critic and a very insightful advisor — He was an undaunted researcher and very much a great spokesman for more work of this type and of things such as are going on at this particular conference, so I couldn't pass without at least mentioning Oscar, at this particular time.

Just to give a hint of the future, I think maybe we'll hear more of this as we go on. Since the later 1950s, early 1960s, there has been a mass of development in multiperson games. Now, in the multiperson games persons may not be individuals, they might not even be groups of individuals, but persons may be incidents and they may have broader applications for operations research military problems than one would think. To give you an example, recently at Cornell University we introduced the Watts telephone system, which is supposed to save about 30% of the telephone bill (I hear they also did this at Yale Univeristy), but they charge everyone the same thing as it cost before the Watts system and put the 30% in the general kitty. At Cornell, the trustees say that you're supposed to break even with telephone service, so the question arose, "how do you now charge for telephone calls if they don't cost as much as they did before?" Well, it turns out that you actually have the axioms for something called the Shapley value, which is treating telephone calls, the incident of a telephone call, as a person. Now, there are a great number and variety of different types of telephone calls, so we approximated this as a game with a continuum of players where calls were players, and then the Shapley value became what people agreed was equity in that situation, and, therefore, now we have implemented telephone rates. We had to do some numerical integration and do some sloppy work to get back from these nice models. There's a book by Robert Aumann and Lloyd Shapley on values on nonatomic games, coming out of the Rand Corporation that sounds like a military book, but it's purely mathematical and one for which you would not expect to find too many applications. Yet I can imagine all types of applications in the regulatory agencies, in terms of various ways of assessing costs and so forth.

One area of game theory, which over the last dozen years has experienced a great advance that has not been given full notice and that is probably more important for this group here, is in the direction of repeated games, recursive games and in particular repeated games of incomplete information. Just within the last couple of years, in fact within the last six months, published theorems have appeared that are really three star and four star types of theorems. Very important theorems have occurred in recursive gaming in just the last few years or so. This is a technique which in more specialized forms is used in some gaming models. There are also at least a few developments in bargaining solutions, and the people who are interested in multiobjective types of criteria really should be looking and be aware of some of these new developments in this area. I think maybe Jim Mayberry is going to say a little more about those so I won't pursue them at this time.

This afternoon we are going to go on then, and look at a couple of specific procedures for solving multistage games that have actually been employed. So the next talk is by Zachary Lansdowne on DYGAM, which is such a procedure for solving multistage games.

23 — *Solution Procedures for Multistage Games at the Theater-Level: DYGAM, an Algorithm for Solving Multistage Games*

DR. ZACHARY LANSDOWNE
Control Analysis Corporation

A dynamic programming technique for obtaining approximate solutions

Dr. Lansdowne: This research was sponsored by the Office of Studies and Analysis of the U.S. Air Force, and the DYGAM algorithm was developed by Control Analysis Corporation a few years ago. This presentation is going to be different from your other presentations because I'm actually going to get in there and discuss the details of this computation algorithm.

This chart (Slide 23-1) summarizes the key features of DYGAM, and I'll be discussing these features in more detail in a few minutes. DYGAM is an acronym which stands for Dynamic Game Solver, and it's a general computer package which solves a broad class of multistage decision problems. In particular, it solves the standard deterministic dynamic programming problem, the stochastic dynamic programming problem, and of particular interest to this group, the multistage game.

Slide 23-1 — DYGAM (DYnamic GAMes Solver)

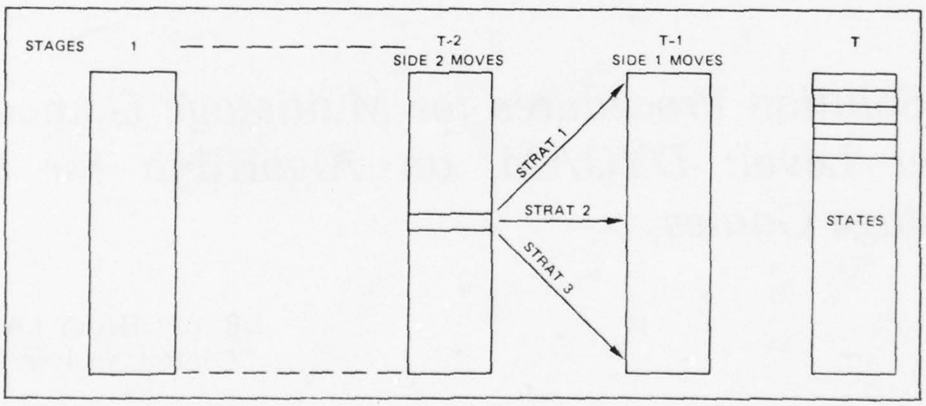
1. Multi stage decision problems formulated as dynamic programs
2. FORTRAN
3. Payoff functions represented by approximating polynomials
4. High state space dimensions
5. "Look ahead" forward algorithm
6. Master and user routines
7. Transformations to change form of payoff approximating function
8. Accuracy control
   a) State space grid
   b) Extra terms
9. Error analysis

The DYGAM system is written in FORTRAN and it uses the standard that Richard Bellman developed, a backwards recursive formula of constructing payoff surfaces starting from the final stage to the next final stage, and so on. The main problem in dynamic programming is the so-called curse of dimensionality. That is, a lot of calculations have to be done, and the way DYGAM tries to speed up these calculations is by approximating the payoff surfaces with approximating polynomials. But because we're introducing errors due to the fact that we're only using approximate polynominals, DYGAM tries to compensate for these errors by using a forward look-ahead algorithm, which I'll talk about in a few minutes.

The DYGAM system has two sets of routines: a master set and a user set. The master set of routines remains the same from one application to another, while the user set of routines changes. If you want a land battle, then you write one special set of user routines; if you want an air battle you write another set.

The user is given various ways of controlling the accuracy of these approximating polynomials, and we have a procedure for estimating the error which is introduced in these approximations.

SLIDE 23-2 TWO PERSON GAME, SEQUENTIAL MOVES



This slide (Slide 23-2) illustrates the basic game which DYGAM tries to solve. DYGAM distinguishes between two classes of multistage games. One game we call the sequential move game, the other game is the simultaneous move game. The sequential move game is like chess or checkers. First, during one stage one side moves, and the next stage the next side moves, and back and forth. The simultaneous-move game is one in which both sides move at the same time during the same stage. DYGAM can handle both, but the simpler case is the sequential move game, and since I only have 25 minutes I'm going to just discuss that particular one; that is, a game in which only one side moves at each stage, more like chess or checkers.

This particular game covers a fixed number of stages, capital T stages. We assume that we have specified as input data sort of the payoff function at the end of the last stage, and this payoff function gives the payoff to, let's say, the blue side, as a function of the number and location of remaining forces for the red and blue sides. We assume that the corresponding payoff to the red side is minus the payoff to the blue side, so we have a zero sum game.

The way a dynamic programming algorithm works is that we start off with a payoff for the last stage, and then we want to solve a series of single-stage problems corresponding to the T-1 stage, corresponding to different

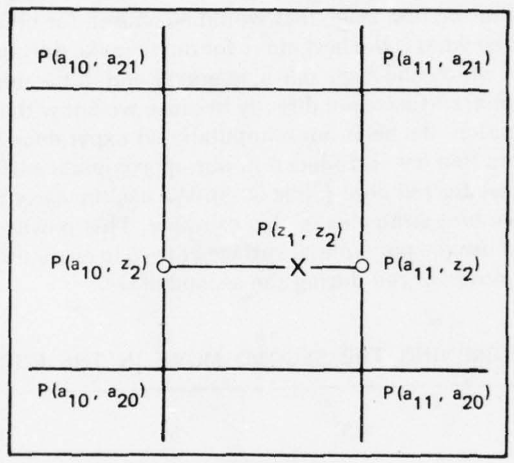
states at that particular stage. The idea in constructing the payoff surface is to solve a series of single stage games for the T-2 stage, and then construct the payoff surface at the T-2 stage.

Now, what DYGAM does to speed up this dynamic programming process is to approximate these payoff surfaces with polynomials. The purpose of this game is to determine the best strategies for red and blue. That is, determine the optimal sequence of blue strategies so that blue will be able to maximize his payoff at the end of the last stage, and the optimal sequence of red moves for the red side, so that red will be able to maximize his payoff.

There are two basic ways of approximating these payoff surfaces. One is to develop a polynomial of very high dimension, which is defined over the entire state space, and the other to divide the state space into smaller regions and define a lower order polynomial for each of these subregions. We tried both ways, and it's our conclusion that the subregional approximation is the better way and that's the way that it's actually implemented in DYGAM.

This slide (Slide 23-3) illustrates the way the regions are broken up. The way our polynomials are defined is that we assume that we have evaluated the payoff surface at each of the corners, or each of the grid points on the diagram, and then we have a formula which computes, or estimates what the value of the payoff is at some intermediate point, as the slide illustrates.

SLIDE 23-3 REGIONAL BREAKDOWN

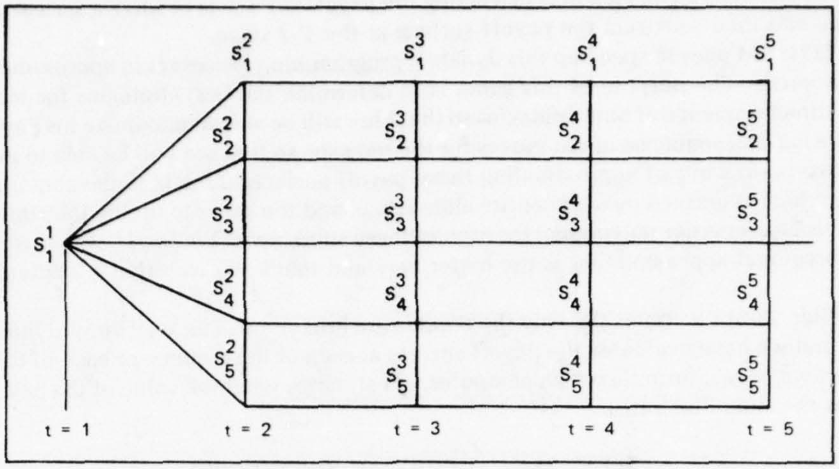


We assume that we have to estimate the payoff at each of the four corner points, and then we have sort of an interpolation function, which enables us to compute the value of the function at the interior point.

Now, if we actually worked with the exact payoff surfaces, then we would be following dynamic programming in its rigorous form and would get truly the rigorous optimal solution. But we're only working with approximate surfaces so, of course, we're introducing error into our calculations. The way we've discovered how to eliminate some of this error is through forward evaluation. Now, what motivated us to come up with the forward evaluator can be illustrated by the same analogy of the experienced chess player which was used yesterday afternoon. The notion is that an experienced chess player who, let's say, is trying to make his first move, will have sort of an intuitive evaluation function which enables him to eliminate most of his potential moves at his first move so that he can search forward in depth, maybe for ten or fifteen moves in the future, on the few remaining potential moves. Then he has another intuitive implicit evaluation function which enables him to evaluate the relative worth of the different configurations that result, ten or fifteen moves ahead, and that enables him to choose the best strategy for his first move.

Well, we used the same idea here. We recognize that our approximate polynomials are only approximations and so we don't want to use them directly to determine the best red and blue strategies. Instead, we only use them as guide posts for this forward evaluation, so we use the early polynomials to eliminate most of the potential strategies. In this case, this diagram (Slide 23-4) illustrates the problem of determining the best move for blue during his first move. We use the approximate surface defined at $t = 2$ to eliminate all but the five most promising strategies. We then use the other polynomials to enable us to search forward in depth, in this case a depth of four stages, to find out what the configurations are corresponding to these five possible moves during the first stage, what the resulting configurations would be at the end of the fifth stage. Then we'd use the polynomial surface to evaluate the relative merits of the different configurations; that is, to find for the fifth stage the five different configurations for $t = 5$ and that would determine the best move way back in stage one.

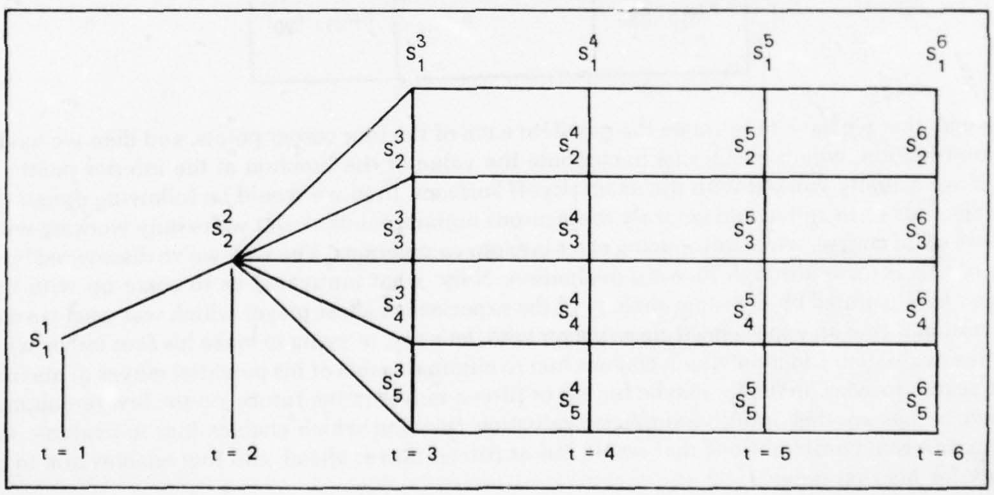
SLIDE 23-4 DETERMINING THE FIRST MOVE IN THE FORWARD DIRECTION



Suppose, for the sake of this example, the best configuration in stage five corresponded to the second branch. That would mean that way back in the first stage that would be chosen for blue so blue would make his second move. Now the problem is to decide what is the best move for red to make during the second stage. Again, I'm just looking at the sequential game in which one stage one side moves and in the next stage the other side moves. So, in using the approximate polynomial surfaces, not directly because we know that there is some error in them, but as guide posts for forward evaluation, it's been our computational experience that this forward evaluator does compensate for some of the errors that are introduced in our approximate surfaces.

Again, we do the same process for red here (Slide 23-5). We use the early polynomial to weed out the poor strategies and retain only the five best strategies in this example. This means we're able to search forward in depth, in this case four stages, and use the polynomial surface at t = 6 to evaluate the resulting configurations and, in this way, determine the best move for red during the second stage.

SLIDE 23-5 DETERMINING THE SECOND MOVE IN THE FORWARD DIRECTION



If these polynomials were correct, were actually true payoff surfaces, then that would imply that the payoff corresponding to these several points, S_2^3, S_2^4, S_2^5, etc. (Slide 23-5) would all be the same. But the fact that they're different is an indication that there is some error in the surfaces, and this observation enables us to estimate what this error buildup might be.

If in Slide 23-6 we define S_k^t to be the state at time t for the k^{th} path along our forward evaluation, and X_k^t to be the corresponding payoff, then if our payoff surfaces were the actual payoff surfaces that would imply that X_k^{t+1} would be the same as X_k^t. But the fact that they're not indicates that there is some error buildup and we define that error buildup as ϵ_k^t. If we make the assumption that the distribution of ϵ_k^t is independent of k and t, we can then estimate the mean and variance of this error buildup. One of the applications of this error analysis is in branch elimination, in speeding up the progress of our forward evaluator. The notion is that it may not be

necessary to go through all four stages for all five potential strategies. It may be possible to cut off some of these branches early, because we know with very high likelihood that a particular branch will not be the winner when we go all the way to the sixth stage in depth, and we use a hypothesis testing procedure to cut off some of these

SLIDE 23-6 ERROR ANALYSIS

S_k^t = State at time t for kth path

$X_k^t = P^{(t)}(S_k^t) \equiv$ Payoff

$X_k^{t+1} = X_k^t + \epsilon_k^t$

Assumption. Distribution of ϵ_k^t is independent of k and t.

Estimate mean and variance of ϵ_k^t.

branches (Slide 7). X_1^v, let us say, is the payoff corresponding to the best payoff at the v^{th} stage, and X_k^v corresponds to the corresponding payoff for the k^{th} branch. The sigma is the standard deviation of our error buildup statistic, and $t + d - v$ is the number of stages we go. The idea is that if this statistic is large enough then that would justify cutting off the k^{th} branch early.

SLIDE 23-7 BRANCH ELIMINATION

Test statistic $\dfrac{X_1^v - X_k^v}{\sigma(2t + 2d - 2v)^{1/2}}$ against Normal.

$t + d - v \equiv$ No. of stages to go.

Now, I don't want to get bogged down in detail, so let me just quickly summarize the main ideas of the DYGAM algorithm. We start off with the basic Bellman dynamic programming recursion and construct these optimal payoff surfaces. But, to avoid the curse of dimensionality, instead of using the optimal payoff surfaces we approximate them with these approximating polynomials. However, to compensate for some of the error that we introduced in these polynomial surfaces we use this forward evaluation technique, and to speed up the forward evaluation technique we have a branch elimination idea (Slide 23-7), and these represent all the basic ideas in the DYGAM algorithm.

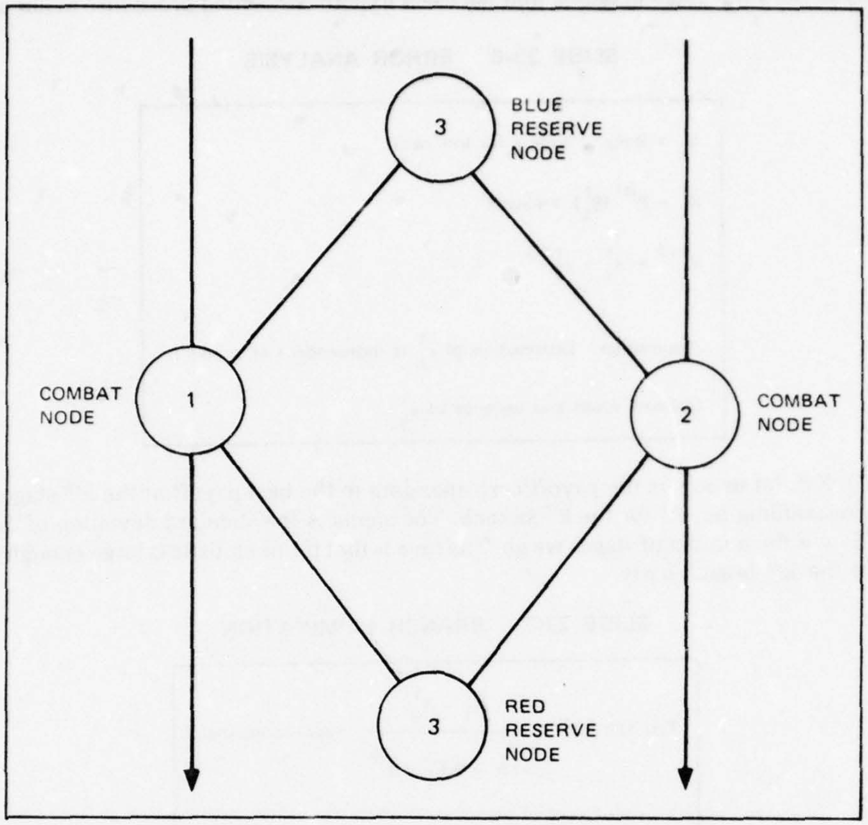
Next, I'd like to discuss some of our computational experience with DYGAM, the running time on a multistage game problem, and some of the areas of current research in DYGAM.

One of the games we looked at is the "two-piston" game, which consists of four nodes (Slide 23-8). Both red and blue forces can occupy the first combat node. Both red and blue forces can also occupy the second combat node. Only blue forces can occupy blue's reserve node, and only red forces can occupy red's reserve node. The assumption is that, when both blue and red occupy the same combat node at the same time, attrition will occur and that this combat node will move up and down on a parallel track. If there are more red forces at this combat node than blue forces, then the red forces will push the combat node back into the blue territory, and similarly for the second combat node.

Now, this particular game can be represented with eight stage variables. We need to know the number of blue and red forces at the first combat node, the number of blue and red forces at the second combat node, the number of blue forces at blue's reserve, the number of red forces at red's reserve, and the positions of the two combat nodes. That is an eight-state variable problem, and that is a large number for dynamic programming. Now, it turns out if you make some assumptions regarding the form of your input data and the form of your final payoff function for the final stage, it is possible to reduce the effective number of state variables to only six. So, basically, this is a problem with six state variables.

Okay, so we solved this particular game with DYGAM, and this Slide (23-9) gives some of our results. We solved the problem with four stages to it, which meant that there are two opportunities for the blue side to move

SLIDE 23-8 THE TWO-PISTON GAME



and two opportunities for the red side to move, since this is a sequential move game, and this particular slide (Slide 23-9) illustrates the results from having two different polynomial approximation surfaces.

SLIDE 23-9 TIMING, ERROR, AND DROP STATISTICS OF TWO NONLINEAR APPROXIMATIONS

	$(g = 3, h_1 = 0, h_2 = 10, h_3 = 33.3)$						$(g = 3, h_1 = 0, h_2 = 20, h_3 = 50)$					
	Backwards Solution	Forward Evaluator (w, d)					Backwards Solution	Forward Evaluator (w, d)				
		(1, 1)	(12, 4)	(6, 4)	(6, 2)	(2, 4)		(1, 1)	(12, 4)	(6, 4)	(6, 2)	(2, 4)
CPU Time (min ±5%)	13.7	0.6	2.8	2.4	1.9	1.6	14.4	0.6	3.5	2.5	1.9	1.6
Mean and		-0.4	-0.6	-0.9	-0.8	-0.9		1.3	0.6	0.6	0.5	0.5
Varience of Payoff Differences		20.7	14.9	14.9	16.3	15.5		36.0	25.9	28.4	28.6	27.9
Fraction of Remaining Paths Dropped at Depths:		1	0.53	0.33	0.39	0.09			0.27	0.17	0.25	0.04
(0.10 level of significance)		2	0.42	0.25	NA	0.14			0.36	0.20	NA	0.02
		3	0.32	0.26	NA	0.21			0.19	0.10	NA	0.11

This particular surface here (on the left) has three grid points for each state variable and the three grid points are at 0, 10 and 33.3. Now, the thing to observe here is that it took 13.7 minutes of CPU time to accomplish the backward solution procedure, the procedure which constructed the approximate polynomials. Then this row here (indicating) gives the various amounts of time required for the forward evaluation technique where we tried out

different forward evaluators. The forward evaluator is defined by two different parameters, the width parameter, which tells us the number of strategies that we've retained for searching forward in depth; and, the depth parameter, which tells us, how deep we actually do our searching. So depending on the width and depth, we have a varying amount of calculations to do in our forward evaluations, but basically it takes about 15 minutes of CPU time to solve this rather simple game which is four stages and has six state variables.

The other rows here, give the mean and the variance of the error buildup, and the bottom row discusses the effects of our branch elimination in speeding up our forward evaluator. The general conclusion is that the two-piston game is a rather simplistic game, and yet it took 15 minutes just to solve four stages of it. So it appears that even though we feel that DYGAM has extended the range for which dynamic programming can be used to solve multistage games, it doesn't appear that we can solve very realistic games in a reasonable amount of time. So there is a need for some more research and more effort to improve DYGAM.

Dr. John Tomlin, of the Institute for Advanced Computation, has looked at ways of improving DYGAM, making it more efficient. In particular, he has looked at the use of parallel processing, and it turns out that there is almost an over-abundance of opportunity to use parallel processing in this algorithm. The basic thing that one wants to do in dynamic programming is to solve a series of single stage game problems that are associated with the different possible states at each stage of the problem. Let's say we're in T−2 and we have to solve a whole series of single stage problems corresponding to the different possible states that the system might be in at T−2. But the notion is that there is no reason why we have to solve one of these states before another state; these can be done in parallel. Furthermore, in evaluating one of these states, we try out different strategies for that state and find out what the payoff is for each one. Again, there is no reason why we have to do one strategy before another strategy; we can do those in parallel.

Then, when we're evaluating the payoff for a particular strategy that is associated with a particular state what we have to do is to evaluate one of these approximate polynomial functions, and it turns out that a lot of the arithmetic that goes into evaluating one of these approximate polynomial functions can also be done in parallel. In fact, this appears to be the best application of parallel processing. Evaluating these approximate surfaces is equivalent really to doing multiple linear interpolation, and a lot of that arithmetic can be done in parallel. Furthermore, in our forward evaluation technique, in which we searched forward in depth on the best strategies, a lot of that work can also be done in parallel.

So, in summary, there are many opportunities to use parallel processing, and to the extent that we're able to do a lot of the calculations with parallel processing, we, of course, can reduce the total running time. What we'd like to do is to try out these ideas on the Illiac 4 computer, which is located at the Ames Research Center in California; it has sixty-four parallel processors. We think that there is a lot of promise and potential in using parallel processing in DYGAM, and, in this way, perhaps we can solve reasonable realistic problems in a reasonable amount of computer time.

Prof. Lucas: Would anyone like to ask any questions at this time?

Dr. Bracken: Zachary, does not the forward evaluation's consideration of only a subset of the nodes destroy the guarantee of optimality?

Dr. Lansdowne: That's correct.

Dr. Bracken: Would it be possible with parallel processing either to have the forward evaluation look at the whole network, or to dispense with the forward evaluation and use only the backward evaluation?

Dr. Lansdowne: Anything is possible. We use the backward evaluation to develop these polynomials, and we try to weed out the poorest candidates. If we didn't use any backward evaluation at all and just stuck with a real complicated forward algorithm, the only way to guarantee optimality is to completely enumerate everything. Now, if you want to drop the forward evaluation and just do the backward. . .

Dr. Bracken: No, I meant after the backward evaluation you do the forward evaluation with a part of the network.

Dr. Lansdowne: Right.

Dr. Bracken: I'm suggesting that you then do it with the whole network.

Dr. Lansdowne: Well, that would be kind of exhaustive enumeration of all the nodes. What makes the forward evaluation reasonable and not reduce itself to complete exhaustive enumeration is that we have some confidence in our backward approximate surfaces. We don't have complete confidence that we'd want to use the surfaces by themselves to determine the strategy, but we have sufficient confidence that they can kind of get us in the ballpark of what the most reasonable strategies are. So all these approximate surfaces are used for is just to try and get us in the ballpark and then we use a forward algorithm to evaluate that. Am I addressing your question?

Dr. Bracken: In the absence of bounds, one's confidence cannot be proven.

Dr. Lansdowne: That's right. We don't have a guarantee for optimality. That's clear.

Dr. Bracken: Do you think it might be possible to just use the backwards evaluation with the parallel processing and get reasonable answers?

Dr. Lansdowne: Right, the only way to guarantee optimality is not to use any approximate payoff surface at all and, instead, to actually determine the exact payoff surface at each stage. However, the user can control the inaccuracy in that the user has the power of choosing the grid size, the number of grid points for each dimension, and so it really reduces to the fact that you might think what we're doing is really taking a nonlinear function and approximating it with a piecewise linear function. Now, if we make the grid size for our piecewise linear functions fine enough, our piecewise linear curve will slowly approximate, the nonlinear irregular curve. And if you want to make this grid size finer and finer and finer, you will be approaching in the limit the real curve. Because we're using the rigorous dynamic programming procedure as the basis for our algorithm, we should, with the rigorous curve, be getting the rigorous answer. So the user does have some control of the error that he's dealing with by determining the grid size. But the only way to guarantee optimality would be to actually use the real payoff surface at each stage and not use an approximation. It's our experience that by using these approximate surfaces, not directly but as guide posts for the forward algorithm, we're able to eliminate most of the worst strategies. This means that when we use the forward algorithm we're just starting in the ballpark, and we just have to evaluate the most promising candidates. So our experience numerically is that the forward evaluation gives pretty good answers even when we have very crude approximating polynomials. But to actually guarantee optimality we'd have to have the rigorous dynamic programming, with optiminal payoff surfaces at each stage. Okay? Question?

Question: Do you have a rigorously optimal solution in this two-piston game?

Dr. Lansdowne: Yes, yes, we did do that.

Question: I didn't see that on the slide.

Dr. Lansdowne: It wasn't illustrated on our slide, but we did extensive hand calculations, and to compare our different surfaces and determine the validity of our different forward evaluations, we actually did come up with the true optimal solution. In our report we do show slides of the results.

Question: Zachary, when you compute these surfaces that you talked about, do you compute the terminal one first?

Dr. Lansdowne: Right.

Question: And then the next one . . .?

Dr. Lansdowne: Well, the terminal one actually is given to us as input data. As part of the parameters of the study, we specify what the final payoff function is for the last stage, and that is kind of the goal function.

Question: Okay, but, as in dynamic programming, if you must work backwards, you cannot compute the k minus the first one without having computed the . . .

Dr. Lansdowne: That's correct.

Question: Do you still use the parallel processing?

Dr. Lansdowne: No, we don't use parallel processing in the sense that we solve T—2 at the same time we solve T—1. As you're pointing out, we still work with stage T in a sequential way. We then do T—1 and then we do T—2 and then we do T—3 and so forth, working backwards. There's no parallel processing here, but within T—2, we're focusing only on T—2 and what we have to do is evaluate a whole series of states within T—2. We can use parallel processing. For each state we have to evaluate the different strategies that are available and there, too, we can use parallel processing. But, you're right, we go back in a sequenced way in terms of backward recursion.

Prof. Taylor: Zach, you haven't actually run any computational things yet with the parallel processing?

Dr. Lansdowne: That's correct.

Prof. Taylor: It sounds like you've got some very good inputs. You've got current state of the art with the people you're talking to. What are their hopes for how much percent reduction in computation?

Dr. Lansdowne: I wouldn't want to hazard a guess. Dr. Tomlin will be part of our panel. He's done a lot of the research in the application of parallel processing to DYGAM, and he'd be a good man to ask that question.

Question: This is sort of a related question. It seems to me that DYGAM in general, and maybe the parallel processing scheme in particular, could help if T were big and the number of states were small. What about the other way around, if T were small, two or three, and the number of states were to be very big? Is DYGAM useful there, or is it not useful but would be useful with parallel processing? You said there are two ways to make the problem big, T and state space.

Dr. Lansdowne: Right. Well, I think DYGAM works best when we have a lot of time periods and just a few state variables.

Question: But what about the other way around?

Dr. Lansdowne: The other way, well . . . that is generally true of dynamic programming in general. In dynamic programming, you want just a few state variables, but you can run it with a lot of stages. That is just a general fact of life in dynamic programming, and that's better than the other way around.

Prof. Lucas: Our final paper this afternoon is by Frederick Miercort, and it's another solution procedure, a solution procedure used in ATACM.

24 — Solution Procedures for Multistage Games at the Theater Level: Descriptions of the Solution Procedure Used in ATACM

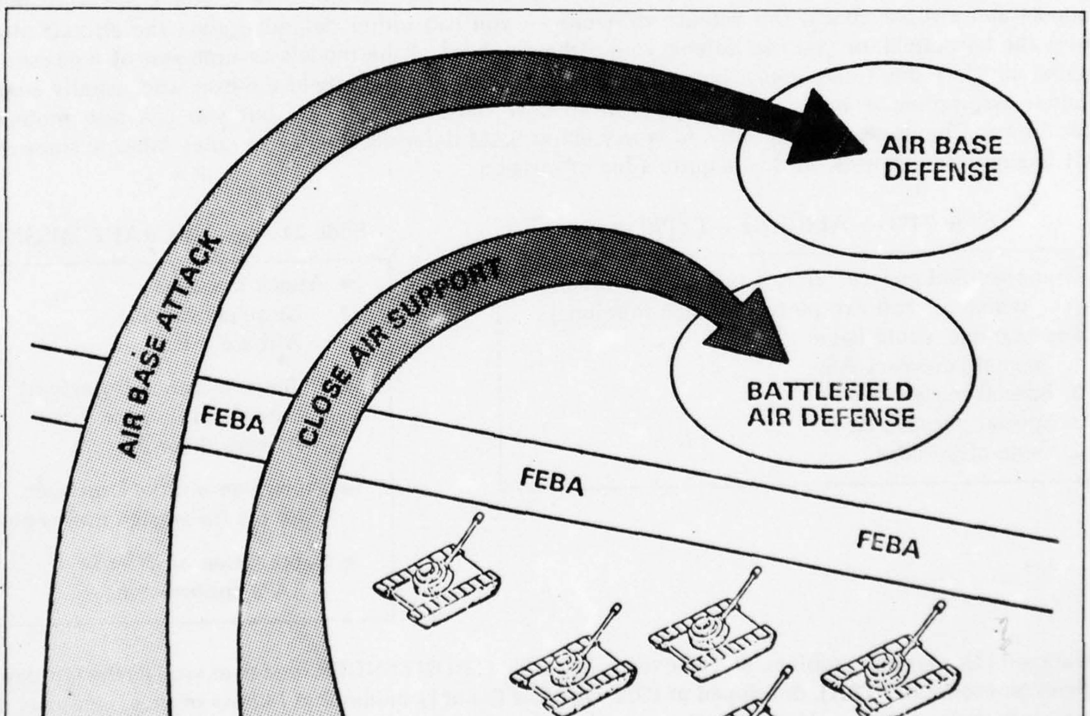
DR. FREDERICK MIERCORT
Consultant

Use of a "conservative play" concept to solve multi-stage games

Dr. Miercort: A couple of years ago when I was at a company called Ketron, Incorporated, we undertook to build for the Arms Control and Disarmament Agency yet another tactical air campaign model, of which there are a number around. I'd just like very briefly to outline the philosophy and the kinds of things we wanted to consider which led us then to the solution approach that we in fact used.

Now, the tactical air campaign — this is really a simple minded picture (Slide 24-1). Basically, what you have is two sides, red and blue, and they both have tactical air forces, which can do several things. You can use them in close air support to support the ground forces; or, you can use them to go somewhat deeper

SLIDE 24-1 PICTORIAL DEPICTION OF PRINCIPAL MISSIONS OF TACTICAL AIR



and attack the other guy's airplanes on his airbases and destroy his airbases. In either one of those two roles he's likely to meet defenders. Now, over the course of some campaign of a specified length, people often talk about a 90-day campaign but it can be any length you want, the object is to try to optimize some overall measure of effectiveness which reflects how much good you're doing to the guys on the ground. It might be total tons of ordnance delivered in support of the ground force, or it might be some measure of FEBA movement — in other words, how far you're pushing the guy back or he's pushing you back.

Well, we did a survey of the available models at that time, and let me just maybe put that off a little bit. We were concerned (Slide 24-2) with a couple of different things; one, the number of different aircraft types that you want to model as opposed to lumping them all together into some supposedly generic or equivalent aircraft type; the number of different missions we wanted to consider; what kind of an objective function might we want to use; and then, given all of the above, how do we go about getting some sort of solution in a finite amount of computer time.

Slide 24-2 — OUTLINE OF RECOMMENDED APPROACH

- Aircraft types
- Missions
- Objective functions
- Methodology

For the purposes to which the model was to be put we wanted to have a fair amount of detail on aircraft types and number of possible missions, which led us then to try to be as efficient as possible in the solution procedure we chose, and that sort of dictated how the thing turned out. Let me just illustrate some of these. For example, you might have some special purpose aircraft types (Slide 24-3), special purpose close air support aircraft, a special purpose intercepter, or a special purpose airbase attacker, and then maybe one or more general purpose types of aircraft that can do any of the above. More specifically, the kinds of missions we were concerned about (Slide 24-4), and it is a fairly large list, included the generic attack missions, the close air support and airbase attack, the defense missions — you can either defend against the aircraft attacking you over the battlefield, or you can defend your airbases. A lot of the models assume sort of a generic intercept mission. They don't distinguish between airbase defense and battlefield defense, and usually some sort of implicit assumption is made in the model as to how those are split up, but you can also protect your aircraft. You can have escorts, you have to worry about SAM defenses, and so you may allocate some of your aircraft to suppress defenses, so this is quite a list of missions.

Slide 24-3 — AIRCRAFT TYPES

- User-specified generic aircraft types and missions (i.e., which aircraft can perform which missions). For example, could have:
  1. Special purpose CAS
  2. Special purpose INT
  3. Special purpose ABA
  4. General purpose

Slide 24-4 — AIRCRAFT MISSIONS

- Attack missions
  - Close air support
  - Airbase attack
- Defense missions (intercept)
  - Battlefield defense
  - Airbase defense
- Protection of attack aircraft
  - Escort (to engage interceptor)
- Suppression of defense
  - SAM suppression

Bearing on this type of problem, we have the original TAC CONTENDER model, as well as the Optimal Sortie Allocation Model (OPTSA), developed at IDA, which is the only model that I know of that rigorously solves the multistage game problem. However, the running time of OPTSA gets completely out of hand if you go to more than three stages. So, what you typically do is to take a 90-day campaign and, say, divide it up into three increments so that your strategy, the way you employ your forces, is the same over that entire 20 or 30 days in each of those stages. Also, the model can only handle three generic missions of close air support, air base attack, and in-

tercept. So we have one general purpose aircraft type, three missions, three stages, and to get an optimal solution, you crank away for a good fraction of an hour on a high-speed computer. Okay, that sort of outlines the problem.

Well, in using ATACM, it has been applied to situations (Slide 24-5) in which for example, we had a close air support aircraft, and we really have two employment choices for this aircraft — it can either actually be used for close air support or we may choose, if the situation gets difficult enough, not to fly the aircraft on a given day so there's a "nothing" mission which says you don't fly them. The intercept aircraft can be divided into either of the two intercept roles. The airbase attack aircraft can either do airbase attack or choose to do nothing. Finally, the general purpose aircraft may do any one of the eight things shown in Slide 24-5.

Slide 24-5 — SAMPLE MISSIONS

Aircraft	Missions
CAS	CAS, nothing
INT	ABD, BD
ABA	ABA, nothing
General Purpose	CAS, ABA, BD, ABD, CAS escort, ABA escort, forward def. supp., rear def. supp.

So, we have sort of eight vectors to the general purpose aircraft, and then two vectors for the other guys. So that problem is significantly larger in its dimensionality than say the problem that OPTSA or TAC CONTENDER considers.

In addition, we wanted to model up to a 90-day war, which would be a 90-stage game. Now, some of the measures of effectiveness are the following (Slide 24-6) — and you can use any of these you want, or a weighted combination of them. One measure might be the differential in tons of ordnance delivered, or it might be some measure of firepower — the difference in firepower. You can put in a FEBA movement model and then the payoff might be the total FEBA movement where you're trying to push them back as much as possible. Or you can use some combination of these things.

Slide 24-6 — MEASURES OF EFFECTIVENESS

- FEBA movement
- Cumulative BLUE minus RED ground plus air firepower
- Cumulative BLUE minus RED air firepower
- Cumulative BLUE minus RED tons delivered in close air-support (CAS) mission

Now, basically what we're talking about is a staged game (Slide 24-7), just as the previous speaker talked about. We have a sequence, say, up to 90 or some such number, of zero sum two-person games. Blue is trying to maximize the sum of some payoff function and red is trying to minimize it. Now, each payoff function is the func-

Slide 24-7 — STAGED GAMES

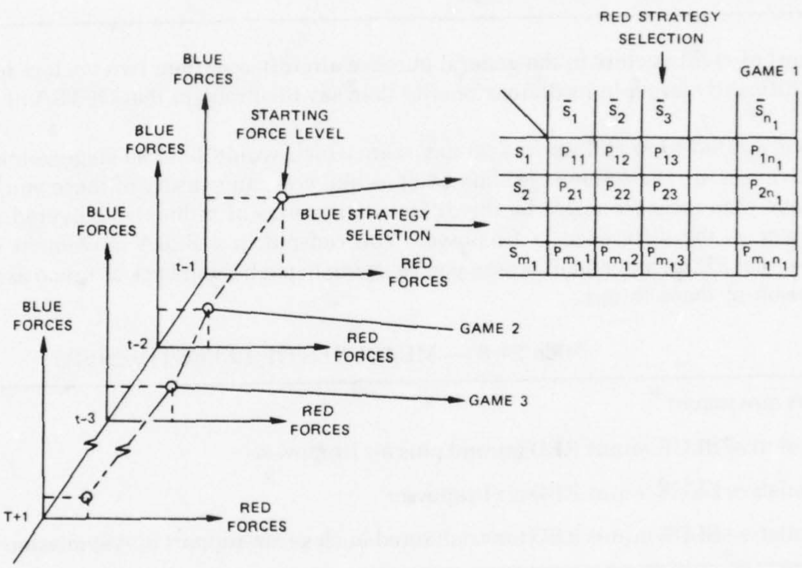
- Sequential, zero-sum, two-person games.
  - — Two players engage in a sequence of two-person, zero-sum games.
  - — BLUE wishes to maximize the sum of the individual payoff functions while RED acts to minimize this sum.
  - — Each payoff function is a function of the resources available to each player at the beginning of each game in addition to the strategies employed by each player during the game.
  - — At the end of each play, new resources are calculated depending on the resources available and the strategies played during the game just finished.
- Attrition relationships, drawdown equations
- Payoff functions
- Methods of solution

tion of the resources available to you on that day of the war and how you choose to employ those forces. There at the end of each day your forces have suffered some attrition, you have leftover forces, and tomorrow you have to decide how you're going to employ your survivors and you play a new game. This continues for the entire scenario.

Now, some very important factors are the attrition or drawdown relationships that you employ in the model, but for my talk today they are really not important. They may be important to the number you get and the answer you get but as far as the model's concerned, you can unplug one set of assumptions or one particular assumption about drawdown equations and plug in another one. The model couldn't care less. But these assumptions are very important in the credibility of the answer you get.

Again, the payoff function can drive the kind of an answer you get, but as far as the solution approach is concerned, it's not really of any consequence. Now, the method of solution that we chose to employ is reminiscent of DYGAM in a certain sense. Let me just illustrate in the figure (Slide 24-8) what this stage game thing is all about.

SLIDE 24-8 A STAGED GAME



Imagine, if you will, that you have a set of times and we have Blue forces and Red forces, pretending like we just have one resource right now, and you start out at some initial value and you select a strategy. The strategy basically says what fraction of your force are you going to apply to the various missions it can prosecute, and both guys do that and that's where the game comes in. So, all the strategies in terms of how you can apply your forces to the missions are on one side, and his (the other side's) strategies are on the other side. This is a very large game by the way. Then, with your drawdown and attrition equations, calculate the payoff. The payoff turns out to be the number of tons of ordnance you delivered versus what he delivered or whatever. You also calculate how many resources you have left. So, the next day, you come down to the time $t = 2$, and you have some small amount of resources. Now you start the whole thing again. We assume here that both sides know how many resources each guy has, and then they play a game that day, get some more payoffs, and go finally to the end of the war. This also assumes they both agree about when the end of the war occurs, which may not be too reasonable. Then you sum up your contributions to the payoff over the 90 days and that's the measure of how well you did.

Well, the solution procedure we developed is based on a couple of things (Slide 24-9). One, we wanted to look at a finite state space. Mathematically speaking, you could look at a continuum, a state space where the amount of resource of any given type could be any number up to some maximum. But, we're talking about airplanes here so let's say you have 2,000 airplanes on each side. Your state space really is integer airplanes. There are only 2,000 of them, but the problem would get completely out of hand if you looked at individual airplanes. So typically you divide the state space into a set of discrete points for each of your resource types, and then you also come up with

a finite strategy space. Theoretically there are an infinite number of ways you can split your forces and prosecute these various missions, but we took what you might call a conservative play by both sides. It had two advantages: we thought it sort of made sense in terms of the problems they were trying to address, and also it made it sort of computationally feasible. The conservative play means we assumed that both sides would play so as to minimize the worse that could happen to them if the other guy had essentially perfect information.

Slide 24-9 — BASIC CHARACTERISTCS OF ATACM APPROACH

- Finite state space
- Finite strategy space
- "Conservative" play by both sides
- Use of dynamic programming to obtain solutions to the multistate problem

Now, to this business of making the set of decision vectors finite (Slide 24-10) — decision vector specifies the fraction of your aircraft type that you're going to allocate to these various missions. So, let's say we have a situation with four missions, and we make it finite by defining a minimum allocation fraction — both TAC CONTENDER and OPTSA and other models use the same kind of approach. A minimum allocation fraction says that if you're going to allocate aircraft to this mission at all it has to be in increments of this fraction that you specify — a half or a third or a quarter or whatever you want to use. So, for example, with four missions and an allocation fraction of a half, the set of possible decision vectors is given as shown on Slide 10. You can either allocate zero or a half or all of your resource type to any one of the four missions, and this is definitely a finite number. As the allocation fraction gets smaller, and/or as the number of missions gets bigger, you still have quite a large set of these things, so it's a large strategy space but it's finite.

Slide 24-10 — DECISION VECTORS

- A *decision vector* specifies the fraction of an aircraft type allocated to each possible mission.
- The number of possible decision vectors is made finite by specifying a *minimum allocation fraction*, f.
  - For example, for 4 missions and f = .5, the set of possible decision vectors is:

(1, 0, 0, 0)	(0, .5, 0, .5)
(.5, 0, 0, .5)	(0, .5, .5, 0)
(.5, 0, .5, 0)	(0, 0, 1, 0)
(.5, .5, 0, 0)	(0, 0, .5, .5)
(0, 1, 0, 0)	(0, 0, 0, 1)

Now let me sort of show you what this conservative play business means. Here's a one-stage zero sum game matrix (Slide 24-11). Let's say Blue is over here on the left and its various strategies are listed down the side and Red is at the top. Now, let's say Blue is trying to be conservative, and he says, "well, if I pick strategy S-1 and Red knows that, he'll choose strategy S-2 and the payoff is only 1, and, gee, if I choose strategy S-2 he'll choose either 3 or 4 and the payoff is zero." So, those row mins over there show the worst that can happen to him if Red is very smart and knows what he's going to do. So his conservative approach is to choose the strategy that maximizes those minima. Now, he says, "Well, the way to maximize the worst that can happen to me is to pick strategy 1 and then I guarantee a payoff no worse than 1." Similarly, Red goes the other way around, and he says, "if Blue's very smart and knows what I'm going to do, he can force the payoff up to 6, 5, 4, or 3, respectively, so I'd better choose strategy 4 and limit the worst that can happen to me to 3." Now, if they both in fact play those conservative strategies, the actual outcome would yield a payoff of 2, which is between the two. Well, we took that same idea

and extended it to the multistage game case. I just mentioned that you don't actually have to completely solve the game to get those values. You just get the row mins and the column maxes and zap, you have it. You don't actually have to solve the game, and that speeds things up quite a bit.

SLIDE 24-11 ONE-STAGE GAME MATRIX

RED MINMAX ↓ ($\bar{S}_4$)

BLUE \ RED	$\bar{S}_1$	$\bar{S}_2$	$\bar{S}_3$	$\bar{S}_4$	ROW MIN
BLUE MAXMIN → S_1	2	1	4	2	1
S_2	3	5	0	0	0
S_3	6	-1	4	2	-1
S_4	3	1	2	-2	-2
S_5	0	-3	-5	1	-5
COL MAX	6	5	4	3	

Let me show you an even simpler example here (Slide 24-12). Here, we have three missions: close air support, air base attack, and intercept on both sides. Again, if Red knows what Blue is going to do he can make the payoff as low as 1, 2, or 0, respectively, so Blue would choose strategy 2, which guarantees a payoff of at least 2.

SLIDE 24-12 ILLUSTRATION OF RECOMMENDED METHODOLOGY

BLUE \ RED	CAS	ABA	INT
CAS	4	1	6
ABA	3	5	2
INT	2	0	2

	0.8	0.2	0	
0.4	4	1	6	1
0.6	3	5	2	2 ←
0	2	0	2	0
	4 ↑	5	6	

MAXMIN = 2, MINMAX = 4, EXP. VALUE = 3.4

On the other hand, turning the situation around, Red may be driven up to 4, 5 or 6, so he'll choose strategy 1 which gives him a payoff of 4. If they actually play it the payoff will be 3. Well, you can solve this game and get the actual mixed strategy solution, and I have shown the probabilities attached to the optimal solution. The optimal

solution says that Blue picks strategy 1 with probability 0.4 and strategy 2 with probability 0.6; Red picks strategy 1 with probability 0.8 and strategy 2 with probability 0.2. Now, if you were to repeat this game many times and each time you played it you picked your strategies with those probabilities, the expected payoff, the average payoff, over many plays of the game would be 3.4, and that is the standard game theoretic solution.

Well, a tactical air campaign doesn't happen many times, it's going to happen once and it's hard to think of a commander flipping a coin or grabbing a random number each day and deciding today I'll use close air support, or if the toss turns out differently, deciding to use airbase attack. You like to think in terms of a deterministic strategy that he would actually employ, and if in fact he employs the mixed strategy solution, he runs a certain risk of doing worse than the conservative bound he could have guaranteed, and for the simple problem you can illustrate it by these numbers right here (Slide 24-13). If, in fact, both sides play their optimal strategy, the actual outcomes that can occur, and the attached probabilities of occurrence, are listed in (Slide 24-13). On the other hand, if the max/min outcome is 2, Blue can guarantee a result at least that good; if the min/max outcome is 4, Red can guarantee an outcome no worse than that; if they both play those strategies they get 3, as we saw before. Now, the expected value is 3.4, but in any given play of the game they're each incurring a certain risk if they choose to use the game theoretic solution. For example, the outcome has a probability of 12% of turning out to be 5, which is worse than Red could have insured if he'd used his min/max strategy. So you might say that 0.12 is a measure of Red's risk, and similarly the 0.08 is a measure of Blue's risk.

Slide 24-13 — PROBLEM ILLUSTRATION

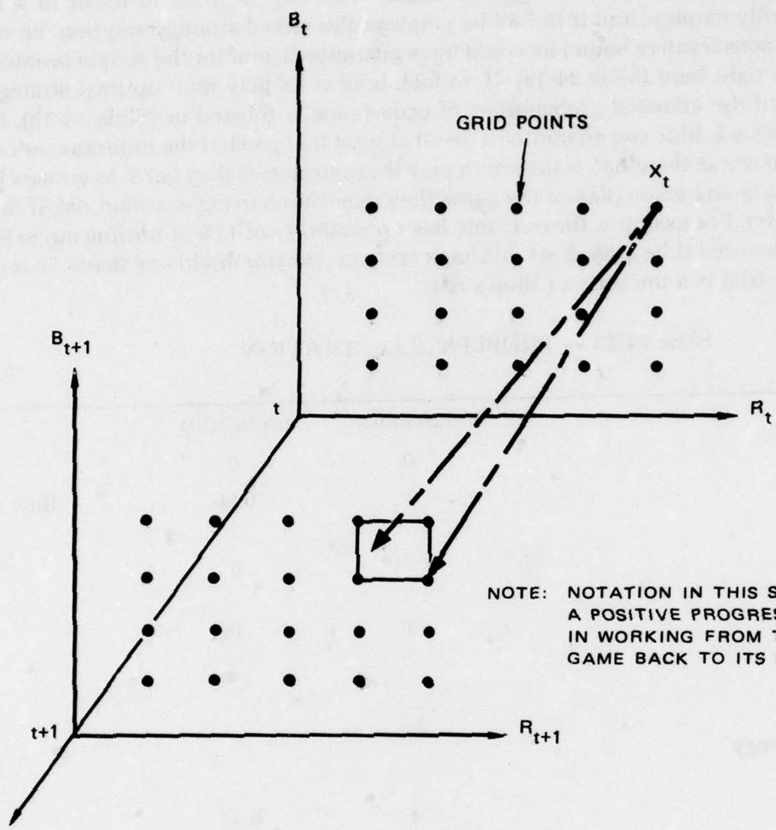
	outcomes	probability	
	0	0	
	1	0.08	Blue's risk
Maxmin	2	0	
Conservative play by both sides	3	.48	
Expected value using optimal mixed strategy	3.4		
Minmax	4	0.32	
	5	0.12	Red's risk
	6	0	

The model implements this buisness of using this conservative play. Now, to illustrate what the model actually does (Slide 24-14). Computationally: it uses a dynamic programming approach, a la DYGAM, and starts at the last day of the war and works back to day one. The model sort of operates in a couple of passes. First, we discretized the state space, and this is represented by the dots for both Red and Blue. Essentially, we solve the following dynamic programming for each player in turn. Let's say we're solving it for Blue; we start at the last day of the game and take one of those points — maybe it's 500 aircraft for me and 400 aircraft for the other guy. Given that I have that, what should I do? I look at all my strategies, assume that the other guy knows what I'm going to do and picks the worst possible thing against me. Then I pick the max/min strategy. Then I remember what that max/min strategy is and tabulate the value of the game, or, the value of that increment. I do it for all the dots in Slide 24-14.

Then I back up to the next to the last day of the war, and I say, "Okay, suppose I have 600 airplanes and the other guy has 500 airplanes." Then I look at the game and where my payoff is and I'll pick a strategy. The other guy is going to figure out what it is and do the worst he can against me and the payoff will be the sum of what I deliver today in terms of bombs, or whatever, plus I already know what my payoff is if I behave optimally on the last day, so I add in the payoff from the last day and I gradually work my way back until finally I get to day one. Now, what do we do? Well, what happens is when I go through my attrition equations I don't end up on one of the

grid points. I end up with 426 aircraft instead of 400 or 500, so to estimate a conservative strategy we then use the DYGAM interpolation approach — we use a linear interpolation in the payoff surface. We actually end up at that in-between point in that little box (Slide 24-14). Well, to estimate the payoff we then use linear interpolation on the results at day t, and sort of make an estimate then of what the payoff is. We use those estimates to get our strategy, so we're using this linear interpolation to get outselves a strategy for all stages of the game.

SLIDE 24-14 DYNAMIC PROGRAMMING APPROXIMATIONS



Let's say we've done that. Now we want to nail down an honest-to-God max/min result. We want to know the worst that can happen to us if we employ this strategy, so now we go back and solve it again. So, if we end up at that indeterminate point we round our resources down and the other fellow's resources up so that we end up at a corner point from which we can get an exact answer. So then you can follow this through the game where you round one guy down and the other guy up at each point and get an actual value for the game, which is a guaranteed mathematical bound to the worst that can happen to you if you employ that conservative strategy. The strategy was developed using the approximation. The bound is developed rounding one guy up and the other guy down. Then, you repeat the whole process for the other guy for his min/max situation, and then, finally, you take the strategy and actually play them together in one forward evaluation to see what the outcome really is. So, the model produces sort of three numbers: the max/min outcome, the min/max outcome, and then sort of the outcome that would occur if they both employ these conservative strategies. That approach has proven to be computationally feasible with several aircraft types, eight or ten missions, and a 90-day war, so it has greatly extended the dimensionality of the problems we're considering. But in this approach you have to buy this sort of conservative approach that we assume would be employed by the two commanders, and that has generated a philosophical debate between us and the guys that did OPTSA and some of these other models.

So, that's what ATACM does, any questions? Sir?

Mr. Asbed: Suppose one of them decided to use CAS, which is antitank, and the other one decided to go airbase attack, which defends the airplanes. How do you put together a payoff that is one number?

Dr. Miercort: Well, your CAS aircraft will, after suffering some attrition, deliver some amount of firepower in

support of the ground troops, and if he chooses airbase attack, he won't get any payoff from bombs on the battlefield. On the other hand, he'll destroy a lot of your aircraft. In addition some of your aircraft will have suffered attrition as they were attacking in the area of the FEBA, so you'll actually get an immediate payoff — which is, you delivered x tons of bombs, and he delivered zero, but he'll attrite your aircraft a lot more than you attrited his. We've already calculated, in the dynamic programming procedure, what would be the optimal payoff from having that set of resources available. We've actually calculated it at a finite set of grid points, but we'll approximate the payoff by using linear interpolation in one of those grid squares. For each stage, we add up the immediate outcome plus the optimal use of our resources from then on, which the backward approach is giving us.

Mr. Asbed: You're kind of finding some kind of equivalent between airplanes and tanks. Are you weighting them?

Dr. Miercort: That's a function of your objective function. Very often people just use tons of bombs. They say I don't know how the ground battle is going to come out, but I can measure tons of bombs, so you measure that. If you want to, you can have some measure of firepower. If you want to, you can stick in a FEBA movement model and use that. The model doesn't really care as long as it has some way to evaluate the payoff function as a function of your aircraft used essentially for CAS each day.

Dr. Huber: Did they check various payoff measures when they ran this model? For example, I'm thinking in terms of the difference model — you called it difference in firepower between Red and Blue, versus a criterion that might use the quotient of Red and Blue firepower, because depending on what the numerical disparities are for the two sides, I suspect you can get quite different results.

Dr. Miercort: Right, we didn't have the resources to really go too much into that. Other people have looked a lot more at that. I think at IDA you guys have done some experimenting with different kinds of measures of effectiveness and how the strategies change and so forth.

Dr. Bracken: Once, I was looking very hard for a set of parameters such that if you used positions of FEBA as a measure of effectiveness you got a different answer than if you used Blue minus Red ordnance delivered. This is in the course of an extended debate, and I was able to find the region that is in the basic paper on OPTSA.

Dr. Huber: I wasn't even talking about a complicated case like that, because you essentially need a land war model to do that, but just look at the quotient of delivered firepower payload versus the difference. I suspect that for a numerically superior side, the difference criterion might be a better payoff than for the numerically inferior, because you may get, for example, an allocation, even though you are producing zero defensive air support capability, because of the difference equation which still permits you to do this. But if you ratioed out the defensive air support potential, you wouldn't do it.

Dr. Bracken: Well, Bruce has thought more about those ratios and differences than I have.

Dr. Miercort: No, this payoff function is a very serious issue and I'm not trying to downgrade it at all, but that just isn't the focus of what I was talking about. You could plug in another payoff function. I'm not saying it's easy to come up with a good one, and the methodology works the same.

Dr. Anderson: Force ratios embracing ratios of firepower are used to determine FEBA movement. Cross out ground power and don't talk about it as FEBA movement, talk about it as cumulative ratios and it's already ATACM, as currently programmed, without changing anything. It can do whatever it does as a measure of effectiveness of Blue divided by Red bombs delivered as opposed to Blue minus Red. Not Blue divided by Red literally, but considering that in a game theory sense, in the sense that you're indifferent between a 5:1 and a 1:5, not that you're indifferent with respect to a number 5 and a number 0.2. I've got a paper on that for whatever that's worth.*

Prof. Taylor: I think what you were bringing up is just the quantification of the objectives that sometimes lead to different strategies, and I think that is something that is not a direct input into the military art of decision making. I think it's an academic pursuit on what we're really talking about. We're talking about decision analysis, not so much a la Ron Howard but a la tactical decision making and modeling thereof in planning.

Mr. Farrell: I'm going to stay in the frame of the topic you really don't want to be on, the payoff function, but a different aspect. The time horizon; I've always been very interested but have never really seen a description of whether it's true that the recommendations about short run strategies are reasonably indifferent to the time horizon once the time horizon gets long enough. That would make me have some belief that short run strategies out of this, or, a true game theoretic solution, or any other solution to this problem, might have meaning. It seems to me that no matter what we've got it's going to be very hard to turn this into real meaning. Can anyone in this field talk to me about the qualitative properties in that aspect of the strategies as a function of horizons?

Prof. Taylor: Well, I've looked at that, and if you look at it analytically even one-sided things become very challenging. I can't speak so much about a two-sided look, but let me tell you about some optimal control

*Bruce Anderson, "Antagonistic Games", IDA Paper 1204 (August 1976)

problems where I looked at analytically determining mathematically optimal solutions.* I found out that it was a function of the parameter values on the payoff, and some of the model parameters, but there was a range where it was very sensitive. When I physically interpreted that particular range of payoffs, as, for example, where there was one type of Blue force that could potentially be directed against two types of Red forces, it became sensitive to the time horizon and the resources available to the players when the Red force with powerful weapons was valued relatively less than the fore with less powerful weapons. In other words, if you value forces and their weapons in proportion to their kill capability against you this would be a case where you violate that principle and put more value, say, on an enemy rifle than on an enemy machine gun.

What I'm saying is that, yes, it is sensitive. It turns out the sensitivity in the one case that I looked at analytically was in a case that didn't make much sense militarily. It does make a difference whether the planning horizon is fixed, or whether the planning horizon is variable depending on various outcomes being achieved. In other words, another thing you can look at is the optimal allocations as a function of the battle termination model, and how you achieve the objective. They turn out to be dependent on that, but again, I did not explore that computationally. I looked at the change in strategies, but I'm not so sure that the payoffs changed significantly.

There was one place where this was looked at many years ago in the TAGS model, and some of the old Rand work, if one can find them in their archives. This was sort of discussed in some of Arnold Mengel's early work, and Lou Wegner there, may remember, or may have worked with it. But Bob (Farrell), you've got an important point there. I don't think that that has been looked at as much as it should.

Prof. Sovereign: Have you looked at the solution in terms of how the allocation vector is changed over a time? What does it look like? Does it seem like it's a reasonable behavior?

Dr. Miercort: Yes, it seems to behave reasonably. At the start of the war you might tend to concentrate more on airbase attack to zap his airplanes, and then when that situation is improved then you switch over more and more to close air support, but as to precisely when and how that happens, or if it does, it can be quite dependent on the relative numbers of aircraft and how effective they are and so forth. We had the contract to develop the model, we gave it to the customer. I don't really know how much they exercised it in terms of looking at various parametric results. They could have done a lot of that if they'd wanted, but I don't know if they really did or not. It would be very interesting to get some insight into some of these things.

Question: DYGAM is capable of doing the same thing, so why not do it with DYGAM? Why is this better than DYGAM?

Dr. Miercort: Well, it's faster. We're not solving the games, we're just picking out the row mins and column maxes.

Question: They're doing that too.

Dr. Miercort: The forward evaluation is very simple and straightforward. We're not exploring all these different paths.

Question: Well, they can do that too, but it strikes me the only difference there is the way they approximate possibly to the way you approximate. Is that right?

Dr. Miercort: Oh, yes, there are a whole lot of similarities.

Prof. Mayberry: I'd like to say a word, in particular, before Fred disappears. In case it should appear to be critical I don't want it to be behind his back if I only say it tomorrow, that it seems to me that this is really very important in this discussion, although you sort of said you're not too interested in the payoff, it's obvious that choosing a good payoff is important. I'm sorry, that's not the primary focus though. But, I'd just like to comment that what is really much more important is choosing two good payoffs, because as someone already suggested, the objective for the one side might be territory, and the objective for the other side might be to preserve his aircraft or some such thing. What this means is that, like any other real problem, to approximate it by a zero sum game is to put it in a very constraining straightjacket, you can't tell how much harm you've done until you look at something resembling a non-zero sum solution.

Dr. Bracken: He does, of course, compute bombs on the results of a solution of a nonzero sum game in that Blue maximizes the minimum he can gain and Red minimizes the maximum he can lose for any measure.

Dr. Huber: One more question; have you ever encountered a case where any of the two sides are able to, or, are being permitted to introduce reserves, that you encounter airbase attack later on in the game but that you didn't in the beginning?

Dr. Miercort: I don't know, I haven't really used the model that much.

Dr. Huber: I suspect, from what I did, that you can always choose the airbase attack in the beginning, and continue doing it with some decreasing effort possibly for some time. Now, if that's the case, you don't need a

*J.G. Taylor, "Recent Developments in the Lanchester Theory of Combat," Preceedings of the Eight IFORS Conference on Operational Research; K.B. Haley (Editor)

dynamic programming algorithm because you could start out at the first cycle, for example, at the beginning, and ask "Well, what happens if Blue and Red select some strategy of airbase attack versus close air support?" What is, let's say the maximum that I decide in delivering payload, close air support payload, and then I solve for the first cycle, and then I go on to the next mission cycle, etc. It seems to me you could have a problem because you have to come up with some kind of a discrete time impulse, and you switch your strategy, for example, after each mission cycle; the mission cycle could be one sortie.

Dr. Miercort: Yes, once you've used the model like this quite a bit and explored how these things behave, if in fact for the first 20 days it is optimal to do something or other and for the next 30 days it is optimal to do something else, you could then go back again to the smaller stage model, say, the three-stage model, and solve it optimally using OPTSA or something similar. But, we didn't want to have any preconceptions about what the optimal strategy was going to look like, and we wanted to give the user at least the possibility of exploring in detail what he should do each and every day. I would think the methodology could also be applied to the nonzero sum case, too, where each guy is using a different measure to evaluate how it's going.

Prof. Mayberry: Jerry Bracken just pointed out to me that in fact you not only can but do get useful results if you simply take two different evaluation functions, because essentially, one guy is going min/max really only caring about his, and the other guy is doing max/min only caring about his, and the question of whether or not those are negative to one another really isn't essentially built into your operation. The only trouble is that if the two objective functions of the two players are quite different then it is true that essentially one player will be defending himself against the kind of attack that the other player isn't the least bit interested in making because he's defending himself against something quite different, so that it might end up being very far from a reasonable pair of strategies but your bounds are still valid, so the bounds that you've got are still good in any case.

Dr. Miercort: Right, any other questions? Thank you very much.

Prof. Lucas: I would like to make one little announcement and then suggest that that wasn't quite the last paper. Gary Brewer of Yale University wanted to participate in this conference, but he had a conflict and couldn't come. However, he wanted me to remind you that he is editor of the *Journal of Simulation and Games* and he's very interested in papers of the type that are being presented here.

Mr. Low: I know it's been a hellishly long day, but Trevor DuPuy is with us today. He's going to have to leave and I think he has a few thoughts he'd like to pass on to us before he goes. He won't be here tomorrow, so if you'll bear with us.

Col. DuPuy: Well, Larry, what I have to say would most appropriately have come either after the previous session on attrition, or would come as scheduled in connection with tomorrow's panel on data. But since I won't be here, I ask your indulgence because I think I have something useful to say, and that I would like to get off my chest. My remarks are stimulated by several things that have been said at this session about data. There were several references during the session that Dr. Farrell chaired. Robbie Robinson also made the statement that there is a need for input data. Dr. Niemeyer remarked that the outcome of battle depends on the input numbers and Mr. Hurford, amongst many other interesting things, said that the real problem is the data.

Now, we all know data are important, but I think that some of the remarks about data this morning can stand a little historical perspective.

Dr. Farrell and some others have commented on the differences or inconsistencies that exist in attrition rates of some models, and about some perhaps surprising similarities. Well, the reason for the similarities, where there are any, is quite simple. Most rates in U.S. models or simulations go back to some rates developed about 1965 by the Research Analysis Corporation, now GRC. I know something about those rates because my company produced the data from which they were prepared. They don't now look anything like the data that we prepared, but, nonetheless, if you look at a document, usually called *The Army Model Study,** published about four or five years ago, you'll find the sad story about how most of the rates, attrition rates, that I think are still the basic attrition rates that are used in most land force models, were developed from these terribly inadequate data. Not that I think the work that we did was good enough, but we only did a few examples and pointed this out to RAC at the time.

Dr. Bracken: Trev, are you talking about Don Mader's paper?

Col. DuPuy: Yes, I think so.

Dr. Bracken: Because that dealt with FEBA movement rates, not attrition rates.

Prof. Taylor: Yes, that's true, but I think it's RAC Report 36.

Col. DuPuy: No, it may not be the Mader piece, but there is a paper in the *Army Models Study* about attrition rates.

Dr. Bracken: Charlie Warner's paper.

Col. DuPuy: Yes, it may have been.

*Review of Selected Army Models (May, 1971)

Dr. Bracken: But it was not the historical study of all of the administrative snafus, it was just a presentation of a scatter diagram.

Col. DuPuy: Oh yes, I recognize that the paper on attrition and firepower in the *Army Models Study* was not definitive. But let me return to the historical work we did which was part of the background of that paper. I think if anyone were to take a look at the original data and then see the way it was massaged and came out in the various RAC models, and then realize that most of the games, models, and simulations now in use by DoD and the services have this as the basis for their present attrition rates, they would be appalled, as I am. This very limited sample in our casualty rate study was not only inadequate in its scope but has been so manipulated that the rates derived from it are often inconsistent with the basic data. This was pointed out to me by a young Major named Evans who was at that time at STAG. He showed how the rates derived from the data often gave results quite different from what the initial data suggested. In other words, I am simply pointing out that there are no data, for all practical purposes, to support the attrition rates currently in use in most models.

Dr. Bracken: After Warner's paper though, Ed Kerlin, who may not argue for himself, but I'll argue for him, went back and redid the calculations which he had done originally from your data in the ATLAS model, and again showed that if you made enough heroic assumptions you could indeed remove the scatter.

Col. DuPuy: That's right. In other words if you massage the data enough, you can get it to say anything you want. But it is no longer real data.

Dr. Bracken: It's an arguable matter. I'm not saying you're wrong, but . . .

Col. DuPuy: My point, Jerry, is that it is ridiculous to take a small sample of some twenty engagements and to presume that from this, through some sort of mathematical sleight of hand, you can come up with rates that represent the real world. I suggest to you that results such as these aren't the real world.

Mr. Farrell: I would like to interject that of those five models that were looked at, at least two of them did not use rates derived from those data at all. They were, in fact, invented from other weapon system performance data with a less historical data base, and I'm perfectly willing to concede that but the agreement among these rates, therefore, is part of what surprised me. I believe that some of the attrition numbers that were in some of the models were derived from the data base you're discussing, by various stages more or less related to it, but some of them I know were completely independent. I can't speak for IDAGAM. As a matter of fact it's one that I can't say anything about, whether in fact those data are related or not.

Co. DuPuy: Well, the significant point that I wish to make, and it's a point that I think is terribly relevant, is that to the best of my knowledge there is no way in which the data which are now being put into our models can be related to real world attrition data.

Now, it is possible, as has already been suggested at this meeting that we do have the results of small unit or individual weapons tests that we can put together in such a way as to provide a reasonable estimate of the attrition certain weapons will create in small units over a short period of time, but the problem comes when one tries to link a number of these together over longer periods of time — and by long period I mean a day or longer. Here is where we get to the point that Rex Goad was referring to: an important aspect of the problem of attrition rates is how fast these individual actions take place. I suspect that they take place much more slowly than most of our models would indicate. I think that this is a major problem and is the reason our models produce attrition rates that are so high that they tax the credibility of some of the decision makers.

I would like to disagree respectfully with one point that Mr. Farrell made when he said that we can get an understanding of war from comparing models. I suggest that we do *not* get an understanding of war from comparing models. We get an understanding of different *interpretations* of war and, from such a comparison, we don't know which one of these may be closest to the real world. This, I think, is dangerous.

We all have a tendency — and when I say we I include myself and the small historical models that we use to try to relate historical experiences to hypothetical future situations — we tend to convince ourselves that our models are reality, or represent reality, when in fact they are not. I believe that the only way we can understand war is to study war, not simulations, and that we should use the simulations as a means to extrapolate to the future some of the individual insights that we get from a study of real war.

Having disagreed with Mr. Farrell on one point, I'd like to agree with him on another, and that is with respect to the concept of clutter. I think it's the clutter, in fact, which makes the progress of real war slower than our uncluttered models will suggest.

Clausewitz expressed this very well when he compared war to trying to walk in water. He said that it is simple, but that it's both difficult and slow. Our failure to allow for this degradation is why models have credibility problems with decision makers who understand — even if only intuitively — what Clausewitz meant when he compared the processes of war to walking in water.

We have been telling ourselves in this conference about the limitations on what our models are supposed to do, and how they aren't supposed to represent reality but are only supposed to be used to make comparisons or for other special purposes. But no matter what we say about this, we all have the subconscious feeling that our models are trying to reflect reality and, therefore, if our models are, in fact, not able to predict, we tend then either

to kid ourselves or to kid the decision maker, or both.

I think that Mr. Farrell's description of the data problem is all too accurate. Since there is no real world data base from which to extract the data, it takes an inordinate amount of time to develop a model data base. This leads then to the situation in which, for each model, the data base is individual, discrete, and subjective. I, naturally, can't disagree with what Mr. Dunnigan said about history as a test and history as the basis for validation. I do want to suggest that the data are available; they merely need to be compiled and organized.

There's a lot more that I could say. For instance, we at HERO are doing a very brief study, a crash study, to try to find out if there is any need for the Army to undertake a project which is popularly known in some elements of TRADOC and DARCOM as "shoot-em-up." This project is to take a field artillery battalion out into the field and for about three to five days have it fire at a surge rate of fire. The reason for this, apparently, is that General Starry, when he commanded the Fifth Corps, was concerned about whether or not our artillery weapons could in fact fire at the rate that our models are projecting, and if so, he feared that there might not be enough ammunition in the stockpiles. To test this, the "shoot-em-up" project was conceived. But somebody had the bright idea that people undoubtedly have fired artillery at such surge rates in past combat. Thus, we are now investigating such historical rates of fire, and finding much data.

So, I suggest to you that there is a lot of data; it just isn't compiled or organized. When it is compiled then I think we will resolve this problem, so eloquently discussed this morning by Mr. Farrell, of different and inconsistent data bases, and the related problem of the time that we are wasting in compiling various individual data bases for individual models.

Now, I've had my say.

Prof. Lucas: In a book on principles of operations research (OR) by Harvey Wagner, he relegates game theory to just one exercise about the min/max theorem. It's a whole book of 1,000 pages or so. In his preface he makes one remark, "In practicing OR we have found that game theory does not contribute any managerial insights to real competitive and cooperative decision making behavior that are not already familiar to church-going poker players who regularly read the Wall Street Journal." Of course, so often around the university, when you talk to people about game theory, they kind of look at you and say, "game theory doesn't really have applications does it? I mean, outside of the large number in the military." So, they have this image of some huge number of highly secretive routinely used techniques, but we know more about that myth, I guess.

The first two papers this morning are going to be on man/machine interactive gaming, and heuristic solutions. I think the size of most models are going to force us away from completely analytic solutions, even though models should teach us something that eventually we try to do analytically. In reality, I think there's going to have to be some man interaction on many of the models. We'll hear a couple of examples this morning. The first paper then is by Alexander Dobieski of TRW.

25 — Man/Machine Interactive Gaming and Heuristic Solutions: Game Theory in Theatre-Level Modeling: Optimal Solutions and Heuristic Solutions

DR. ALEXANDER DOBIESKI
TRW

Interactive man-computer simulation applied to a combat training device. A hardware/software system description

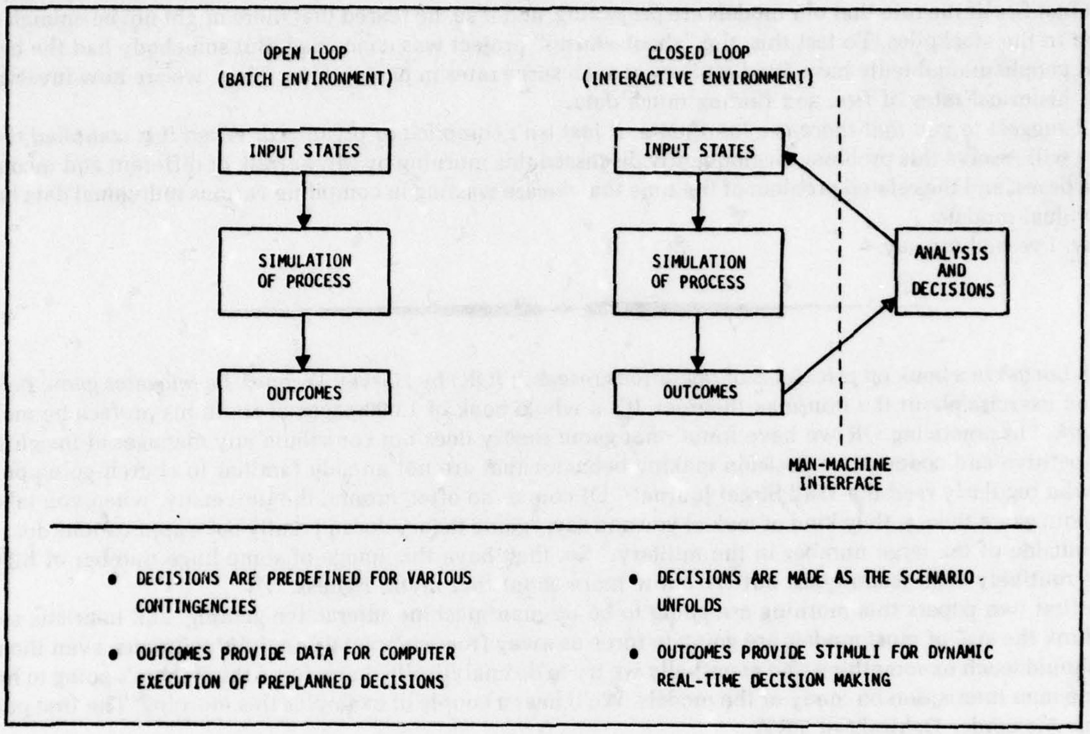
Dr. Dobieski: I want to be consistent with the other speaker's; therefore, I will also deviate from the published topic and will talk on the subject of man/machine interface.

I want to talk about requirements for such a man/machine interface and relate this discussion to some of the problems we've been discussing in this conference. I want to establish the requirements; show you a specific example of a man/machine interface: talk about how it works, what's good about it and what's bad about it; and, where this is going.

So far in this conference we've been talking about modeling, and most of our examples have been open loop or batch environment type models where you have inputs, simulation processes, and outcomes. In this

situation, decisions are predefined and intermediate outcomes are used by the computer itself to make further decisions. Yesterday some suggestions were made that we should move in the direction of having a man in the loop (Slide 25-1). Here, we still have input, simulation, and outcome, but the difference is that decisions are made by the user as the scenario unfolds, and the outcomes provides stimuli for the decisions. Regardless of the kind of modeling we're talking about, we still need a man/machine interface. In the open loop case, the man/machine interface is usually card input and tabular output. In the closed loop case, this interface is usually keyboard input and CRT output.

SLIDE 25-1 TACTICAL COMPUTERIZED GAMES REQUIRE AN INTERACTIVE MAN-MACHINE INTERFACE



Some new requirements are shown in Slide 25-2.

The first requirement (Slide 25-1) is tactical symbols displayed and registered on a map background. This is consistent with what we're talking about in this conference: communicating with people, communicating with the user, getting the user involved. As Dunnigan said, he always has maps with his work. I think whether people want to admit it or not, what they usually do is take the output of a model and mentally if not physically plot the results on a map so they can better picture the situation.

The second requirement is for the easy input of complex decisions — you don't want to require the user to be a computer scientist or learn a whole new language or special acronyms. The third requirement, machine transparency, means that the user can use the device without having to be trained in computer

Slide 25-2 — THE MAN-MACHINE INTERFACE EQUIPMENTS PROVIDE SYSTEM DESIGN CHALLENGES

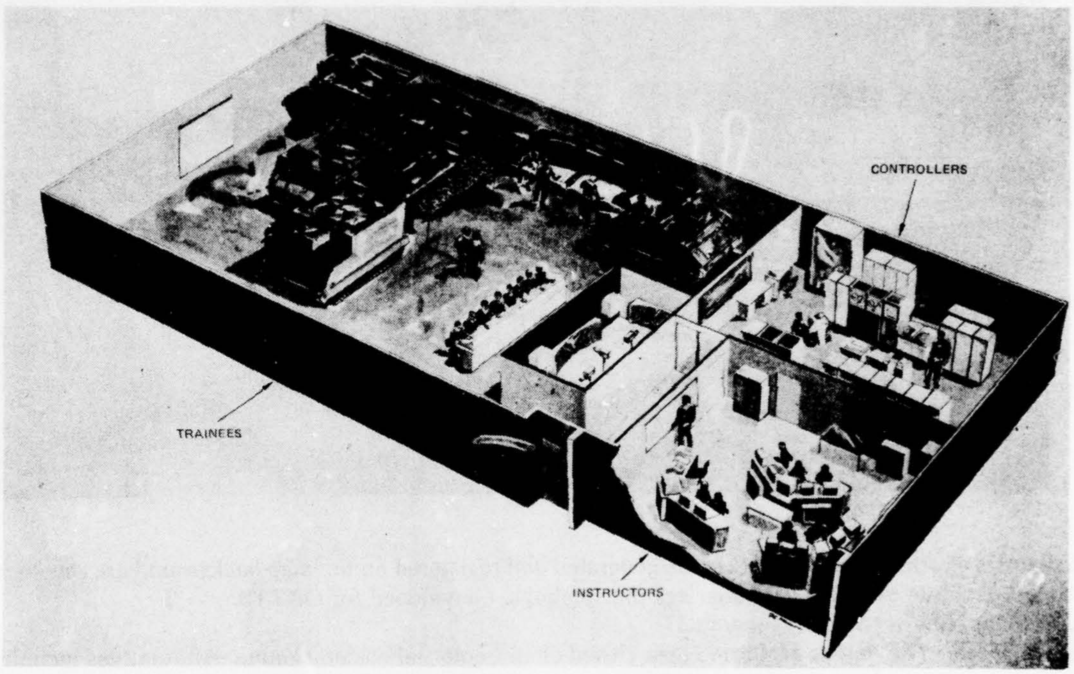
- Tactical symbols displayed and registered on a map background
- Complex decisions easily input to the computer
- Machine transparency
- Minimal training for operation
- Fail-safe operation

CATTS is an example of a system that satisfies these requirements

operations. The fourth requirement, minimal training, is particularly important. A lot of our models take a long time to run. You don't want to require the user to go to school for several weeks to learn how to use the man/machine interface. Finally, we have fail-safe operation — a lot of people sit down to a machine and they're afraid to use it because they're afraid they'll cause a catastrophic failure. The requirement is to design a man/machine interface that is fail-safe (i.e., if the user makes an error he won't ruin the system).

An example of a system that addressed these requirements is CATTS (Combined Arms Tactical Training Simulator), and I'm going to use this example to present a man/machine interface. First, let me explain a little bit about CATTS. I think CATTS is a good example of a marriage between gaming and simulation because it is a true simulation model, it runs without any people at all, or it can run interactively. Yesterday we had comments that we've been in simulation and should move toward gaming. I think this is an example of a true marriage between the two. The system works as follows: People being trained go into a simulated battalion tactical operation center (Slide 25-3) which includes 17 tactical radios, 5 field telephones, maps, grease pencils, the usual equipment that commanders normally have. In CATTS, these trainees talk to people who play the roles of their unit commanders. In the real world, the unit commanders execute the decisions of the commander using troop leading procedures. In CATTS, the role-playing unit commanders execute decisions by inputting them into the computer using the man/machine interface.

SLIDE 25-3 CATTS OVERVIEW

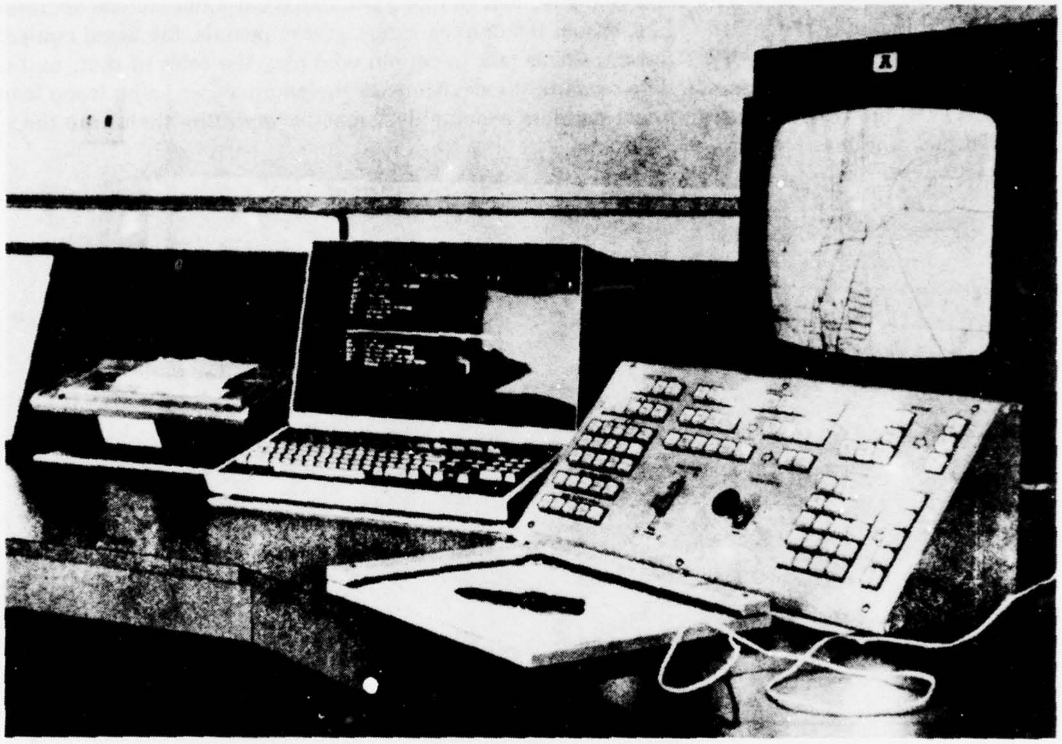


In the real world the units report back. In CATTS, the mathematical model, which runs in real time, reports back to these controllers who report back to their superiors over the simulated communication networks. So, it's a true gaming system in that there is role playing by the controllers. It's a true gaming system in that the trainees (battalion commander and staff) constantly presented with stimuli and they have to make decisions. It's a true simulation in that the game is driven by a real-time simulation model battalion operation, which resides in the computer.

Let's now address the man/machine interface in CATTS. Here there are three controllers controlling the whole game and they do that by operating equipment at the console shown in Slide 25-4. We have an alphanumerical display, hard copy device, color CRT with map background and superimposed tactical symbology, a function box and graf pen. Suppose the controller wants an output. He selects a function button, corresponding to the kind of output he wants, and he gets a graphical picture like that shown in Slide 25-5. They say a picture is worth 1,000 words. I think in this instance it's worth about 10,000 because here you have the relative location of every unit with respect to every other one; you can select the location of the Red units, the direction of movement; symbols that indicate where the weapons are firing — all integrated into a map background. Imagine now for a minute we have this data tabularly. One's mental image of the tactical situation would not be nearly as

clear. The software generates the tactical symbols by driving a color graphics generator, and automatically registers on a map background, so, as the scenario unfolds, the units move according to where they're supposed to be. The user can pan-tilt and zoom on the scene and the units are automatically registered when the map motion stops. Using this output device you can watch the model run. This is part of the visibility we've been talking about in this conference.

SLIDE 25-4 CONTROLLER STATIONS SHOWING (left to right) PRINTER, ALPHA NUMERIC DISPLAY AND KEYBOARD, TABLET WITH STYLUS, GRAPHICS DISPLAY PANEL, AND RASTER CRT



Examples of other symbols that can be generated and registered on the map background are shown in Slides 25-6 and 25-7, while Slide 25-8 summarizes the symbol set developed for CATTS.

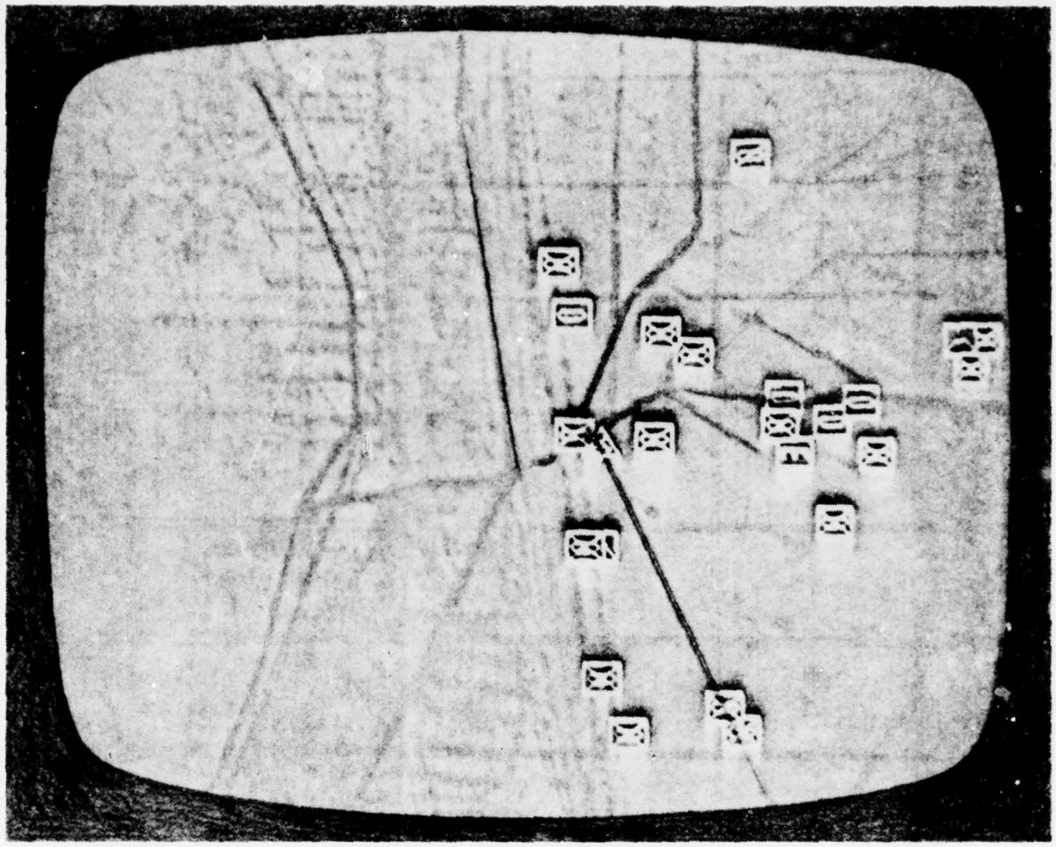
Question: How is the map generated?

Dr. Dobieski: The map is generated by a closed circuit color television camera. Alternatives include a computer-generated map, photographic map storage systems, floppy disc systems, and x/y positioning tables. We used a pan-tilt zoom system with a closed circuit color television camera essentially taking a picture of the map background because at that time it was the easiest and least costly way to satisfy the contract requirements. The user pan-tilts and zooms on the map, and there are digital shaft encoders which go into the computer to tell the computer what image is in the field of view.

So far, I have addressed graphic output. Perhaps the most significant piece of work we did was in the way the user inputs data to the system. The user operates a device called an analog tablet, I'm sure a lot of you have had experience with this, but I'll explain this operation for completeness (Slide 25-9). This particular tablet has acoustic microphones on its edge. The tip of the pen emits an audio tone. When the individual moves the pen around on the tablet the computer calculates the position of the pen and generates a cursor which moves in the corresponding position on the screen. When the operator wants to make a selection, he pushes the tip of the pen down and that signals the computer that is the selection he wants (see Slide 25-9).

To use this device, you first select the proper sequence of buttons, and a menu appears on the bottom third on the monitor. Slide 25-10 illustrates a menu for changing unit length, width, direction, and location. The operator moves his pen over the selection he wants, and makes a selection. The software backlights the selection in red to indicate the item selected. If the operator has made a mistake, he can make another selection. He can continue to

SLIDE 25-5 GRAPHICS DISPLAY EXAMPLE 1



SLIDE 25-6 GRAPHICS DISPLAY EXAMPLE 2

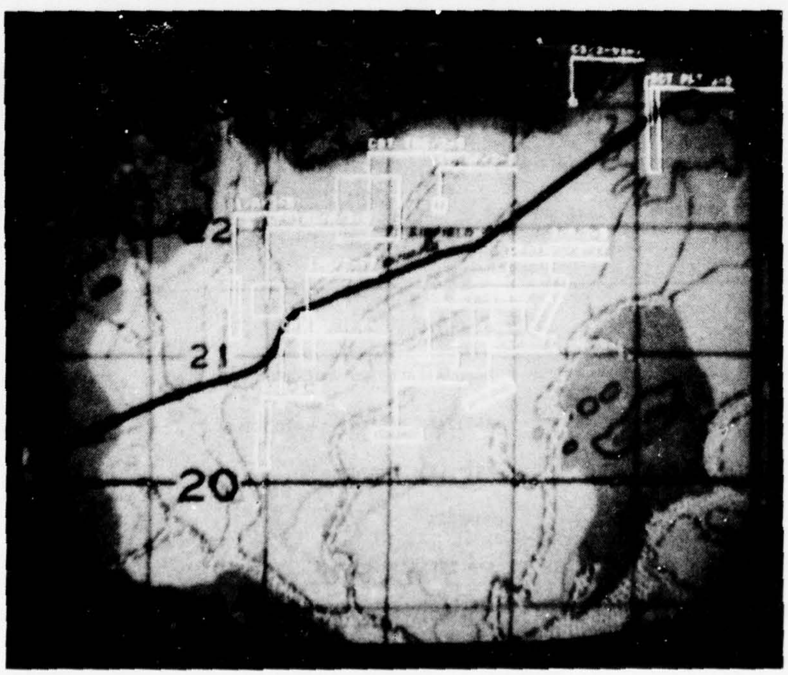
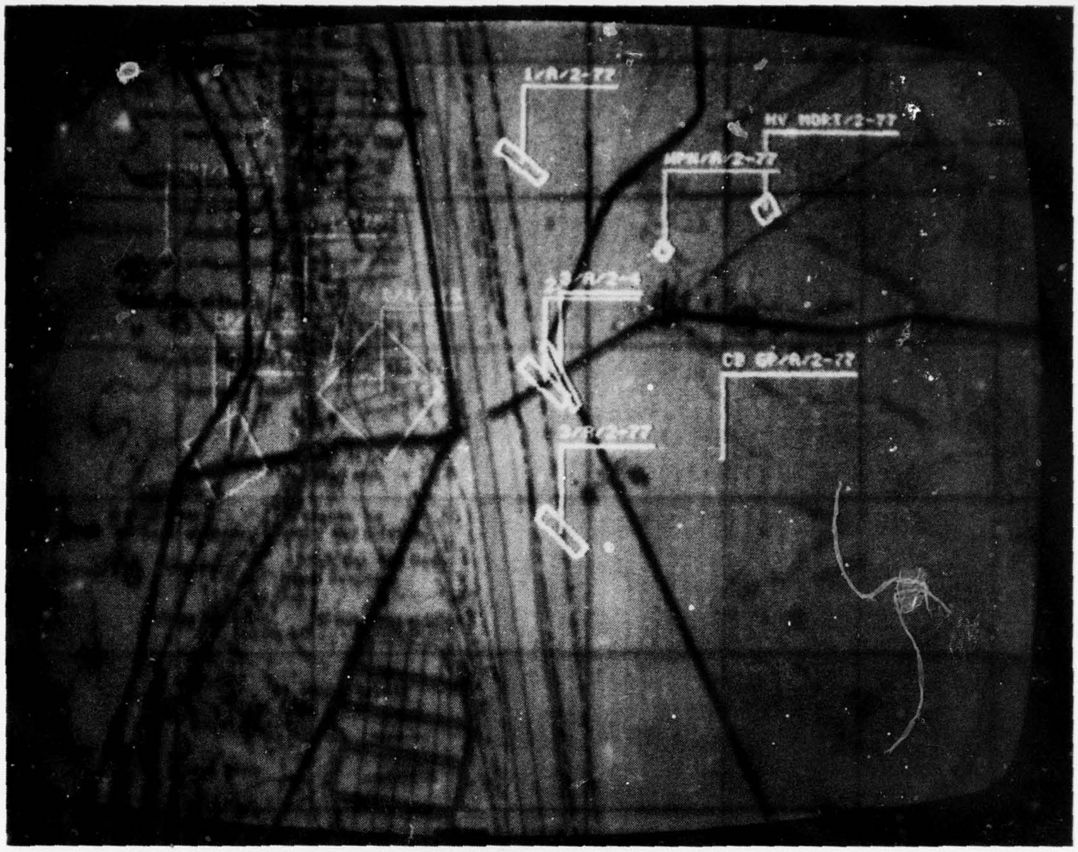


FIGURE 25-7 GRAPHICS DISPLAY EXAMPLE 3

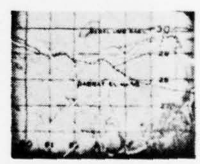


SLIDE 25-8 THE GRAPHIC DISPLAY SUBSYSTEM

PROVIDES FOR RAPID REAL TIME DISPLAY

AND COMPREHENSION OF COMPLEX SITUATIONS

GENERATED BY THE APPLICATIONS SOFTWARE



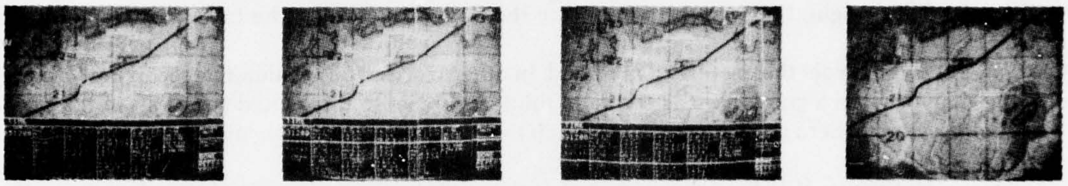
DISPLAY OPTIONS INCLUDE

UNIT AREA OCCUPIED	SENSOR LOCATION AND COVERAGE	IMPACTING FIRES
COMMAND POST LOCATION	ANTITANK ROCKETS AND MISSILES	CONTROL MEASURES
UNIT DIRECTION OF MOVEMENT	ARTILLERY WEAPONS	OBSERVATION POSTS
FRONT LINE TRACE	MORTARS	OBSTACLES
TACTICAL OVERVIEW	AIR DEFENSE WEAPONS	MINEFIELDS
GRID COORDINATES	AIR STRIKES	PREPLANNED TARGETS

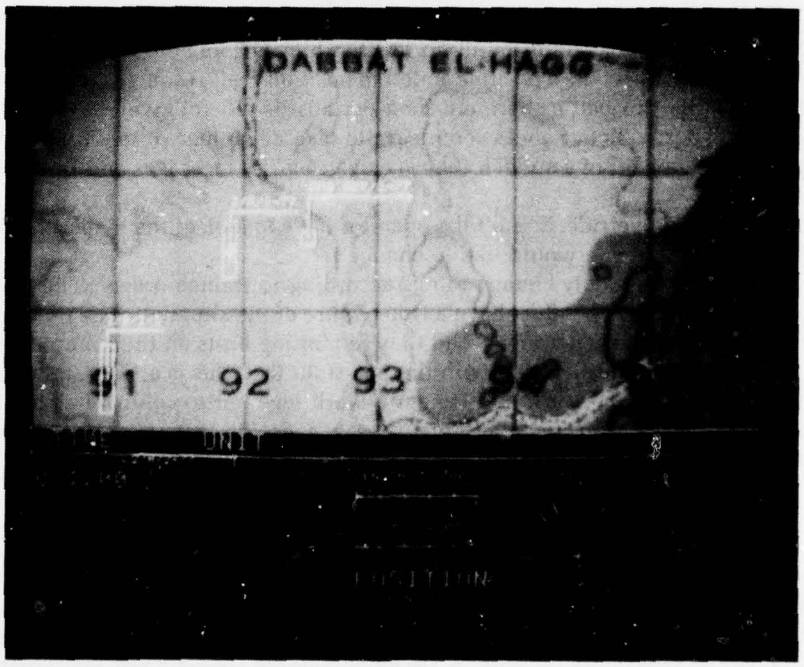
SLIDE 25-9 OPERATION OF THE GRAF PEN AND ACTIVATE UNITS MENU

OPERATION OF THE GRAF PEN

ACTIVATE UNITS MENU



SLIDE 25-10 MEANS FOR CHANGING UNIT DIMENSIONS, LOCATION, DIRECTION OF MOVEMENT



change his selections until he selects "done". This final action changes the data base. In essence, this menu constitutes a high order graphical language. It's syntax free because there is no usable syntax. All inputs are made by pointing at the desired item.

Mr. Asbed: You said this one can run automatically.

Dr. Dobieski: Yes.

Mr. Asbed: What is the time difference in occupying the computers?

Dr. Dobieski: I'm sorry, I don't understand your question.

Mr. Asbed: What is the time difference in getting one run through with and without . . .

Dr. Dobieski: It depends on what kind of scenario you want. The system has run up to 16 hours but is designed to run up to 4 hours. It is a real-time battalion level simulation, so it runs in real time. If you want to simulate 4 hours worth of battalion operation, it takes 4 hours.

Mr. Asbed: To do it automatically?

Dr. Dobieski: Yes, the model runs in real time.

Mr. Asbed: And with the operator . . .?

Dr. Dobieski: Runs in real time also. If the operator doesn't make any decisions, the model is still going to execute whatever the battle plan was. It is up to the trainees to make decisions to control the battle, to change the battle, to modify their original plan based on the situation. It is up to the trainees to get new information and command and control their forces. If they don't, the war still goes on, which is the whole point of CATTS. While people are making decisions things are still happening in the battlefield, and the war is not going to stop for them to make up their minds about what they want to do.

Question: So, you have manual override?

Dr. Dobieski: You have manual override, exactly. Yes sir?

Prof. Taylor: So you're really not compressing time but you're making it sort of independent. This is something the decision maker has to operate with if he's going to have any effect downstream.

Dr. Dobieski: That's right, the point is to duplicate the flow of messages, the traffic loads on those decision makers.

Question: I'm trying to get this picture in my mind. In other words, you wouldn't be receiving messages every minute of the time because a guy who was trying to think and react to this would spend all of his time reading messages. In other words, there is a period of time when there's no stimuli coming out of the machine, is that correct?

Dr. Dobieski: Not really. What is happening is that the stimuli are coming out of the machine as the events occur.

Question: Yes, but I mean there may be a short period of time . . . .

Dr. Dobieski: When there's nothing happening, yes.

Question: And you can make decisions hopefully during some of these times?

Dr. Dobieski: During the build-up — we have a build-up here where the Red units start moving toward the border and what happens is that during this build-up time nothing is going on, if the movement of enemy forces goes undetected. The trainees must take a positive action to position their resources to detect such enemy activity. So, the important point is that in battle there are long periods of lull followed by periods of intense activity followed by periods of lull again. The trainees soon learn to take advantage of the lull periods to do their reconnaissance because, if they don't, all of a sudden they start getting fire messages and find that they are no longer controlling the war.

Mr. Asbed: In this manual override, doesn't the operator have to look at the maps and study them for awhile and make a decision as to where one would like to change it?

Dr. Dobieski: These operators only change things according to trainee orders. The operators have the displays, these trainees do not. So, the operators have a true picture of the situation. The trainee's job is to figure out what the true situation is from message traffic. And they're plotting units on their maps with grease pencils just like they do now, and discussing the situation with their staff. In fact, this is a team training environment, there are 17 people in here all arguing, debating, and trying to work together to control this battle.

Mr. Low: For instance, if they do send out reconnaissance, do the team members get some kind of display that would be commensurate with what you get in real life?

Dr. Dobieski: I don't understand what kind of display you're talking about.

Mr. Low: Well, photographs, for example.

Dr. Dobieski: Not at battalion level, we don't do that, although that's augmentation, it could be done. But they send out a reconnaissance patrol: they just tell it where to go. These controllers input the command using the interactive system and then they can watch the picture. Okay, now this is all recorded for replay later on. So, after the exercise, the whole exercise is replayed for the battalion commanders and certain points are brought up by the instructors, like, "remember two hours into the battle when you did thus-and-so, well, here was the true situation" — that kind of thing. So, it's a training experience, a learning process through experience more or less. There is no attempt to score these people. There is no attempt to tell them they did poorly or they did well. It's up to them to decide how they thought they did. Yes sir?

Question: Have you run experienced combat officers through it?

Dr. Dobieski: The Army has, yes.

Question: What was the reaction then? Did they think it closely approximates or only partially approximates?

Dr. Dobieski: Okay, there have been 13 battalion staffs, probably more than that now, have gone through. The systems have run for two years and, as with new systems, the Army first viewed it with some amount of caution and skepticism. It was controversial at first but now it's being very well received. A commander in Europe said everybody going to Europe will go through CATTS.

Question: Where's it set up?

Dr. Dobieski: At Fort Leavenworth, Kansas. The Army is trying to answer some difficult questions like "does this device train better than less expensive systems?" Meanwhile, the battalion officers coming through are for the most part very enthusiastic about it.

Question: What's the potential for taking these concepts up to the theater level?

Dr. Dobieski: I think it's there, within the state of the art.

Question: Could it be done, for example, by changing your map and some of the positions of your commanders?

Dr. Dobieski: Okay, when you say theater level, I don't know who you're trying to train or who you're trying to get to react. I mean, what level do you want to go to?

Question: Going from a battalion staff to say a theater staff.

Dr. Dobieski: Oh, to a theater staff, sure I think it's perfectly feasible. Yes sir?

Dr. Huber: They're running exercises all the time, but these are exercises where the computers these days are always operated by staff, just like SYNTAG has an annual exercise, or the two allied tactical air forces, and once you look into a staff, you find they do a pretty lousy job in generating noise. I'd rather have a noise generator of this kind because it's probably more realistic than staff where everyone generally has a friend and knows in advance what's going on so he won't be embarrassed during the exercise.

Dr. Dobieski: Okay, one point here, comparing this with a CPX. In the CPX's I've been involved with I often received a message which bore little relationship to what I had been doing. For a lot of the messages I used to get, I'd say to myself "This is impossible, because my forces are over here, so this reported event couldn't happen." In CATTS, there's a reasonable explanation for everything that happens. If you commit one force, fifteen minutes later that force isn't going to be available for something else you want to do. So the interactions, or the effect of your decisions, really come back to haunt you an hour or two hours later (i.e., a non markovian process). I think that is another important thing about this system.

Okay, back to man/machine interface. As an example of changing a command, here are the FORTRAN variables in this particular model to give you movement command (Slide 25-11). Using a conventional interface, the user has to specify all these things, plus there is a software pointer system to point at operational states. Im-

SLIDE 25-11 EXAMPLE — A MOVEMENT COMMAND FROM CATTS MODEL

- **OBJECTIVE**

  COMMAND A UNIT TO MOVE ALONG A SPECIFIC ROUTE

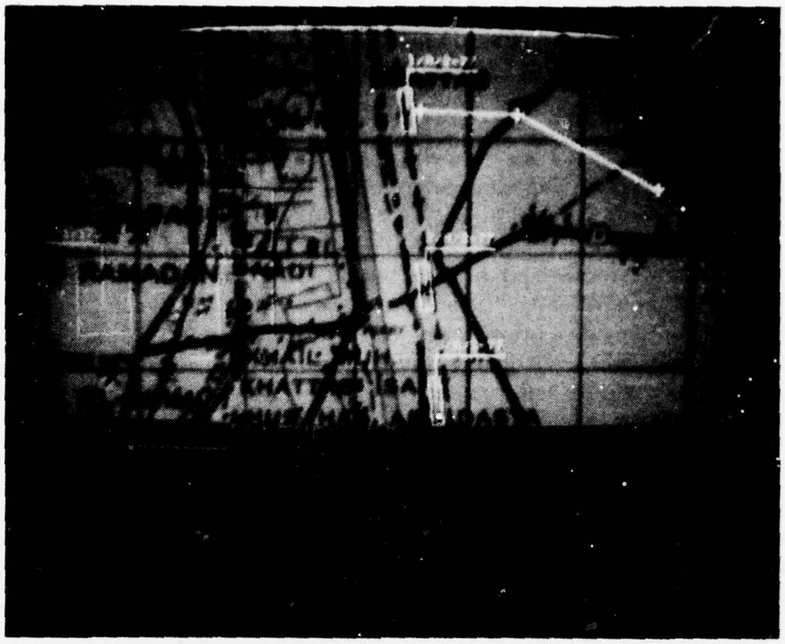
- **IMPLEMENTATION**

  CARD INPUT

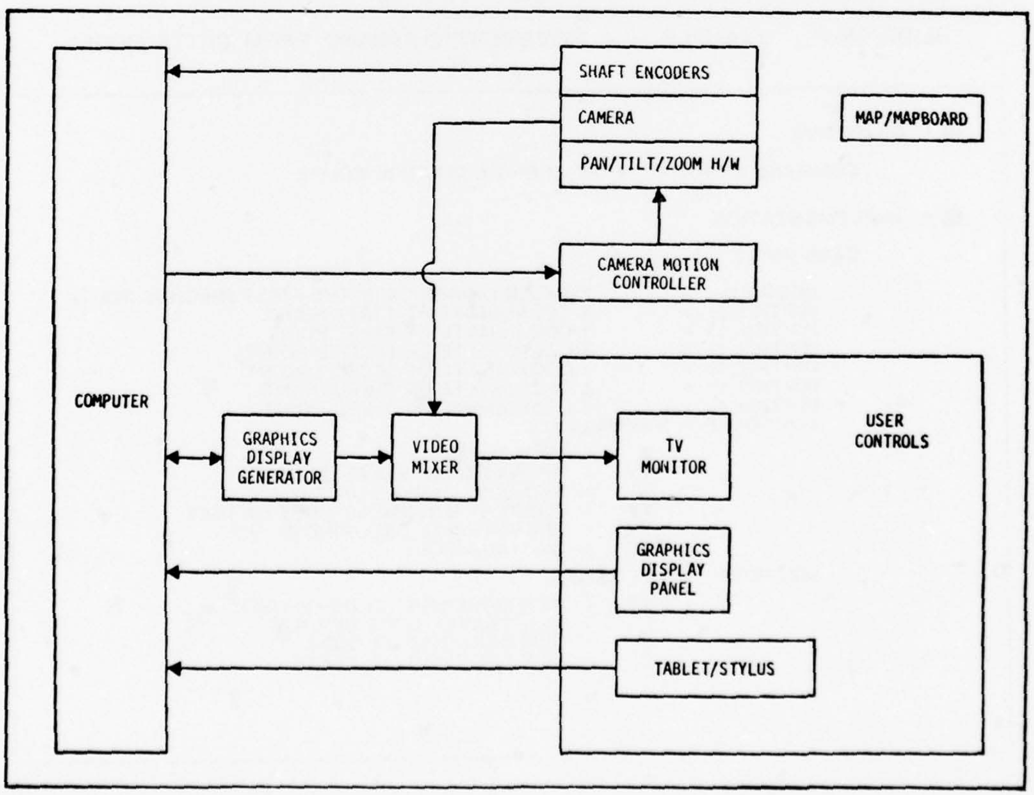
```
NCDPO(1)    =     NUMBER OF POINTS IN THE FIRST SPECIFIC ROUTE
IXPTH(1,1)  =     X COORDINATE OF FIRST POINT
IYPTH(1,1)  =     Y COORDINATE OF FIRST POINT
IXPTH(2,1)  =     X COORDINATE OF SECOND POINT
IYPTH(2,1)  =     Y COORDINATE OF SECOND POINT
IXPTH(3,1)  =     X COORDINATE OF THIRD POINT
IYPTH(3,1)  =     Y COORDINATE OF THIRD POINT
MVCHG1(1)   = VVVWXXYYZZ
                V   = UNIT TYPE
                VV  = OPERATIONAL STATE OF UNIT
                W   = RED UNIT (1) BLUE UNIT (2)
                XX  = CURRENT MOVEMENT CODE OF UNIT
                YY  = OPERATIONAL GROUPING OF UNIT
                ZZ  = UNIT NUMBER
MVCHG2(1)   = EEFGJJKKK
                EE  = NEW MOVEMENT CODE OF UNIT
                F   = NEW TRAVEL CODE OF UNIT
                G   = NEW DEPLOYMENT CODE
                JJ  = NEW OPERATIONAL STATE
                .
                .
                .
```

agine an operator sitting down at a terminal trying to implement these commands by typing in values for the FORTRAN variables, so we obviously need a higher order language. Contrast that with this sytem (Slide 25-12). It

SLIDE 25-12 MOVEMENT COMMANDS



SLIDE 25-13 MAP/GRAPHICS HARDWARE BLOCK DIAGRAM

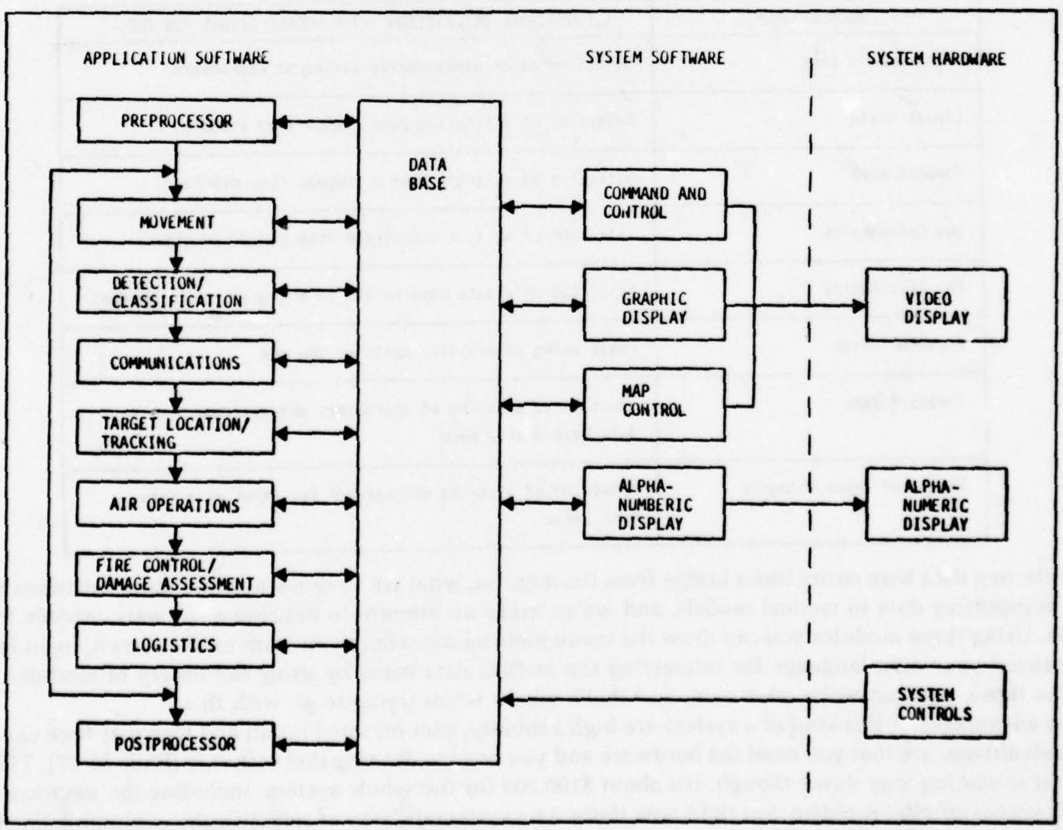


does the same function, exactly the same function graphically. If the operator wants a special route, he actually draws the route on the map with the graf pen, so it is purely map-based interactive graphics. Every selection the operator makes is backlighted in red by the computer to tell the operator what he has selected. So, this is truly a syntax-free, high order language.

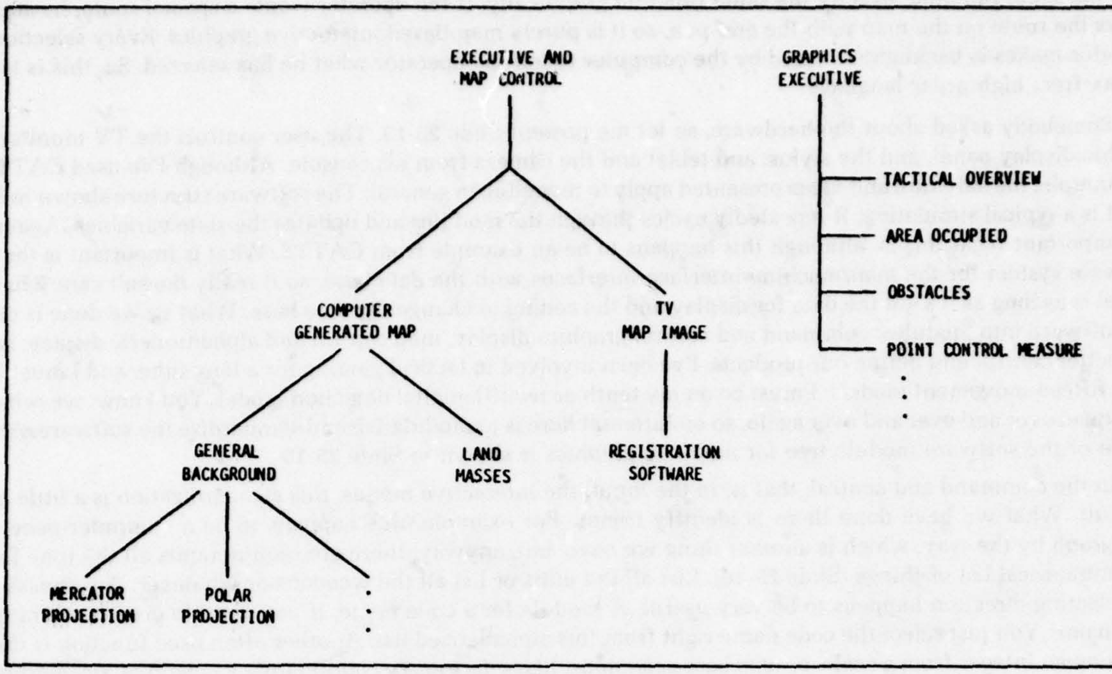
Somebody asked about the hardware, so let me present Slide 25-13. The user controls the TV monitor, the graphic display panel, and the stylus, and tablet and the camera from his console. Although I've used CATTS as an example, the software and ideas presented apply to modeling in general. The software structure shown in Slide 25-14 is a typical simulation. It repeatedly cycles through the modules and updates the state variables. Again, it's not important what this is, although this happens to be an example from CATTS. What is important is that the software system for the man/machine interface interfaces with the data base, so it really doesn't care what the model is as long as it's got the data for display and the coding to change the data base. What we've done is divide the software into modules; command and control graphics display, map control and alphanumeric display, so we can better control and define our products. I've been involved in tactical gaming for a long time, and I must have done fifteen movement models; I must be on my tenth or twelfth digital detection model. You know, we reinvent the wheel over and over and over again, so our attempt here is to modularize and standardize the software. An example of the software module tree for map and graphics is shown in Slide 25-15.

In the command and control, that is, in the input, the interactive menus, this standardization is a little more difficult. What we have done there is identify trends. For example, this happens to be a computer-generated viewgraph by the way, which is another thing we have, but, anyway, there are requirements all the time for an alphanumerical list of things (Slide 25-16). List all the units or list all the weapons or whatever. A compass rose for selecting direction happens to be very useful. A module for a code name, if you want to give an operation a code name. You just select the code name right from this alphabetized list. Another often used function is that of selecting an integer from a scale, or map base selections, like select an xy coordinate, or a chain of xy coordinates

SLIDE 25-14 SOFTWARE OVERVIEW



SLIDE 25-15 MAP AND GRAPHICS SOFTWARE CAN ALSO BE MODULARIZED



SLIDE 25-16 GENERALIZED MODULAR SOFTWARE CAN BE DEVELOPED FOR COMMAND AND CONTROL

MODULE NAME	FUNCTIONAL DESCRIPTION - THE MODULE ALLOWS FOR THE:
Alphanumeric List	Selection of an alphanumeric string of characters
Linear Scale	Selection of a floating point value from a scale
Compass Rose	Selection of an angle from a compass rose display
Map Coordinates	Selection of an X, Y coordinate from a map background
Map Data Entity	Selection of a data base entity (i.e., unit, road, bridge)
Position Units	Positioning of military units on the map
Create A Name	Creation of a string of characters and assignment as a data base entity name
String of Input Integers	Selection of a string of integers for input as a data base value

for a route, or a data base entity like a bridge from the map. So, what we've done is define those functions that are useful in inputting data to tactical models, and we've made an attempt to develop a software module for each function. Using these modules you can draw the menu pictures anywhere you want on the screen, so, in fact, you can construct your own language for interacting the tactical data bases by using the library of modules. If you don't like those, you can write your own, and that's where we're trying to go with this.

The advantages of this kind of a system are high visibility, user oriented input, and true real-time capability. The disadvantages are that you need the hardware and you need to develop that software (Slide 25-17). The hardware cost is coming way down though; it's about $100,000 for the whole system, including the microcomputer. The software is another problem, but right now there is no systematic way of scientifically evaluating the relative

SLIDE 25-17 SUMMARY

GRAPHICS ADVANTAGES

- HIGH VISIBILITY OUTPUT/USER-ORIENTED INPUT/REAL-TIME CAPABILITY/ INCREASED PRODUCT UTILITY

GRAPHICS DISADVANTAGES

- SPECIAL HARDWARE REQUIRED/HIGHER INITIAL DEVELOPMENT COST/ ADDITIONAL SOFTWARE REQUIRED

AREAS FOR FURTHER WORK

- NO SYSTEMATIC WAY OF SCIENTIFICALLY EVALUATING THE RELATIVE MERITS OF ALTERNATIVE DISPLAY TECHNIQUES AND FORMATS
- SOFTWARE HAS BEEN COMPLEX AND COSTLY

merits of different display techniques. For example, nobody knows whether it's better to have blue or green, and human factors people are working on graphic preception problems.

Okay, here is another summary chart, the final chart, showing the advantages and benefits of all this (Slide 25-18). First of all, the user can watch the model run and obtain immediate real-time comprehension. Again, we've been told here by Dunnigan that we have a marketing problem, that we can't interface with the user properly, and that we don't get the user involved. I think this is a good way to do that. I think instead of simplifying our models we can present the results better. I don't think we necessarily have to simplify. It's not that complicated in my mind as long as we present the results properly.

SLIDE 25-18 MAP GRAPHICS FEATURES/BENEFITS

- USER CAN "WATCH" THE MODEL RUN FOR IMMEDIATE AND REAL-TIME COMPREHENSION OF THE TACTICAL SITUATION
- EASY IMPLEMENTATION OF COMPLEX DATA BASE CHANGES BY THE USER
- MINIMAL TRAINING FOR OPERATION
- MACHINE TRANSPARENCY
- FAIL-SAFE OPERATION
- TACTICAL SYMBOLS AUTOMATICALLY REGISTERED ON THE MAP BACKGROUND
- TABLET HAS PEN AND PAPER MOTION - PARALLAX ELIMINATED
- REDUCED FATIGUE AND VISIBILITY PROBLEMS

SYSTEM HAS UTILITY FOR

- SOFTWARE DEBUGGING
- SCENARIO CENTRALIZATION
- MODEL EXERCISE
- POST EXERCISE REPLAY

Easy implementation of a complex data base is another benefit. I've discussed the natural syntax-free language.

Minimal training for operation; people — military officers — who never saw a computer use this system in about two hours and are proficient at it in about two weeks.

Machine transparency; you don't even realize that there's a computer there. Another thing about it is that it has a high amusement content. The high amusement content offers a real opportunity for communications on how the model works, and it's just fun to use. It's like playing Star Wars, or Star Trek or Pong games.

Fail-safe operation; you really can't hurt the system, by pressing the wrong button.

Tactical symbols are automatically registered on the map background. When the model advances the data base changes and the software redraws where everything is.

Finally, we have reduced fatigue and visibility problems. The system has utility for software debugging. You'd be surprised how much easier we're able to debug the model by looking at a picture of it. You can see units moving off in a wrong direction; you can identify data base changes; you can, in fact, display your whole data base pictorally to debug it. Again, when you're setting up your scenario, you can look at a picture of it and debate where you want everybody; it's easy to change things around. During a model exercise, of course, and post-exercise replay, we store everything and then play it back.

So, in summary, I think a man/machine interface system like the one I just presented has a high potential to solve many of the problems we've been addressing in this conference.

Prof. Lucas: We'll take one or two very quick questions now.

Question: I have two related questions; first of all, are there any items that the computer doesn't generate?

Dr. Dobieski: For example?

Question: Well, take a reconnaissance question, or a weather state or something like that. Do the control operators inject any message? In other words, are there aspects of requests from the game staff that may not be handled by the computer so that they can write the messages themselves?

Dr. Dobieski: Yes. The operators are very clever in controlling the tempo of the game, too. Okay, so the man/machine interface, and all that hardware act as an aide to them. I can't give you specific examples, because I can't think of one now.

Question: There is, at least to a degree, some training there.

Dr. Dobieski: Yes, there's some training and experience there.

Question: The second question is, do you do anything in the computer — input misinformation, or fuzz the accuracy to a degree?

Dr. Dobieski: Yes. Sometimes the operators will hold back information or fuzz things up, give a false report, just to see how the battalion staff will react. Again, you don't need a computer for it. Human beings there do that, they control the tempo of the game. For example, there are static and jamming on all the communication nets, so operators can put in static and jamming, say, on the fire direction net and now the players have got to switch nets. Meanwhile, they have all these fire support requests queing up. So, it's very interesting to watch how they solve those kinds of problems in a team environment. The idea is to make the players duplicate the experience that they have in the real world.

Question: You give the impression, quite possibly that in order to apply this to my model, I have to write some software. Does that software reside in my computer or yours?

Dr. Dobieski: Okay, it could reside either place. It depends on what we try to do here. What we're trying to do now is off-load all this graphic software and all this picture software into a microcomputer. This should uncomplicate the host interface. If you don't have a separate microcomputer then you've got to interface all this hardware to your computer and put the software in your computer, and you have to have an operating system that can handle it, so there are advantages and disadvantages both ways, and it can be done both ways. I don't know if I've answered your question.

Mr. Simpson: Thinking of it now, not so much for training, but in the sense of exploring things, how much capability is there now for someone to explore various possible courses of action by trying one and then backing up to a certain point and then trying something else? Is there such capability now, and, if not, how much trouble would it be to put it in?

Dr. Dobieski: Yes the system has been designed to handle this situation within limits.

Question: Like in fire support, suppose a guy wanted to try out one pattern of fire support and then said, "well, let me see now, I want to try another one in which I put more over here."

Dr. Dobieski: Not only that, but you can run the scenario to get a fire support situation, then implement one decision rule, store that, back the system up to where you were, and then try another one and implement that; back the system up to where you were and try a third one, and store them all for replay and comparison. So, you can restart the game any place you want. Suppose something interesting happened, suppose it's hour four and

something really interesting happened in hour three. You might want to back it up to hour three and review with him what he did and then go again. Yes, it has the capability to do that.

Prof. Lucas: Our next paper is by Paul Tuan of SRI International.

26 — Man/Machine Interactive Gaming and Heuristic Solutions: Some Common Problems with Man/Machine Interactive Modeling

DR. PAUL TUAN
SRI International

Some concepts and challenges associated with the development and implementation of interactive gaming techniques

Dr. Tuan: I'd rather listen to Al instead of giving this. Actually, last Tuesday morning, when for the first time I met Al Dobieski, we were talking about topics, because we were afraid, judging from the printed topic on the program, the two talks might be very similar with some redundancy involved. Of course, that fear was very quickly dissipated, and this morning, after listening to Al's very excellent presentation, I feel the two talks can actually complement each other, because his talk is based on some very real experience of the actual system, and my talk will be a little more general and will purposely leave some of the questions vague.

Also, to make sure that we don't have any redundancy, I have an advantage being the second speaker, so when he was here talking, I was very quickly eliminating some of the slides I have with me.

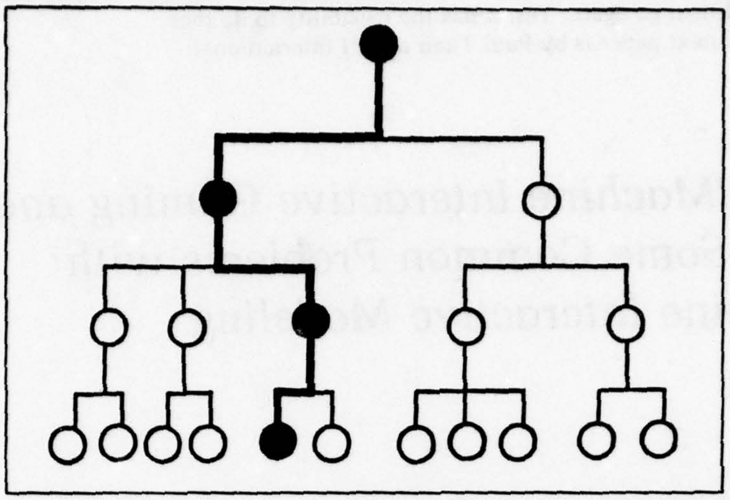
A very short time ago, I probably would not have considered the topic of tactical problems as a suitable one for a conference or workshop of this level, because these problems can be very easily classified as programming problems. But, as we get more and more into man/machine systems work, I am finding out that the problems are quite a bit more than just programming problems, or hardware problems. A very simple analogy can be drawn, for example, that if we didn't have computers today, linear programming or the simplex algorithm would simply remain as a paper tiger. Even with a computer today, linear programming turns out to be a big pain in the neck to run on the machine.

The realization that these problems are more than meet the eye is not only derived from a recent experience in developing an interactive BALFRAM, which by the way is sponsored by ONR with participation from CINCPAC, but also from some other nonmilitary applications. Now, these applications normally deal with very critical resource allocations within a highly competitve environment so some of those experiences directly apply to the situation in hand.

So, what is man/machine interactive modeling or gaming? Simply stated, it is a computer-based simulation or gaming system whereby a person is allowed to interrupt to review the gaming status and to interject his decisions or options on line in a cooperative manner with a computer. Or, as many of you have stated very eloquently in the last few days, it is simply putting a human in the loop. So, I would suspect in the very near future we'll have computer models, man/machine interactive models, instead of naming them ATLAS or IDA or VECTOR 1, 2, 3, probably it will simply be named the GEORGE, George being the first man in the loop.

We can look at this interactive man/machine experiment in many ways. One of the ways we can look at a computer simulation or gaming experiment is simply a progression along a decision tree like this (Slide 26-1), overly simplified, or course, whereby we have many many possible outcomes of the simulation, and in any one simulation experiment we're resorting to one of the outcomes if the problem does not crash in the middle. What outcome it will produce will depend on the particular path that particular simulation will take, and of course, the path is determined by the branching function at each of the decision points.

SLIDE 26-1   SINGLE PATH vs. MULTIPLE OUTCOMES

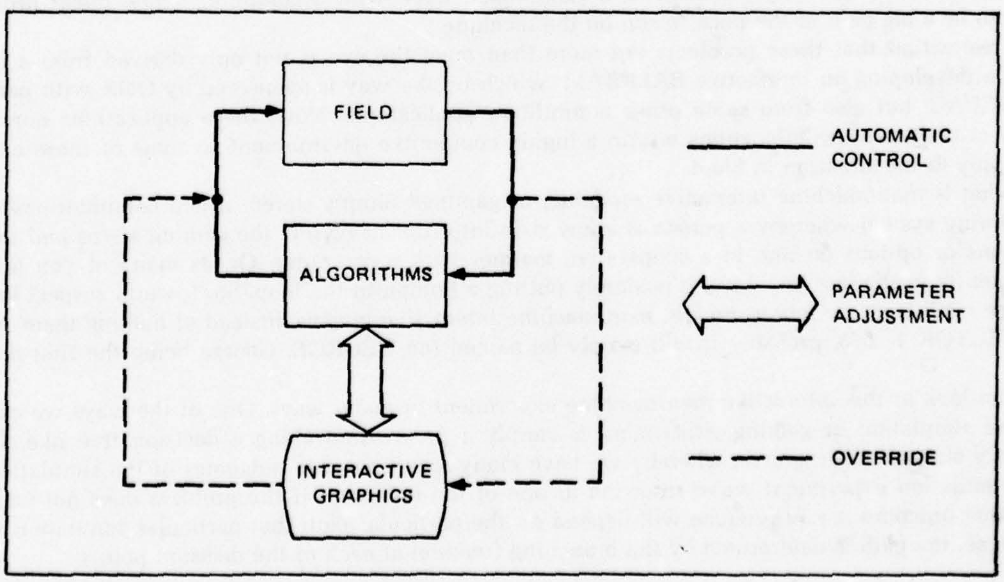


Now, of course, before we run the simulation we have no way to know what path that particular simulation will take, therefore, as computer model users we have to supply all the possible contingency logic data and the decision parameters and algorithms to every possible decision mode that we can think of. This will, of course, result in a horrendous amount of work in designing input and inputting the computer at the outset of the experiment, and because the task is so great many times people try to simplify the input. For example, we know that the attrition — the threshold value of the attrition — where a certain unit becomes combat "not effective" is a percentage. This percentage is maybe a function of the condition of the battle, but to simplify that, we simply say okay, the Blue ground unit threshold is 35% throughout the game, everywhere under any condition it is 35% and things like that.

Okay, for the man/machine interaction then, it is not necessary to supply all the logic at once. In other words, you can actually wait until the simulation reaches a particular point and at that time the human probably would supply the decisions to keep the game going. So this means we'll save some of the data preparation and also give some realism to the simulation.

Another way to look at this, actually, is to take an example from a non-military system. The one I have in mind is a highly critical traffic control real-time system. The system was designed on a closed loop, automatic algorithm and controls the assignment of, say, vehicles. (Slide 26-2). This sytem actually cost the government two million dollars to develop. Not by us, but by a number of people, and the system ex-

SLIDE 26-2   MANAGEMENT BY EXCEPTION

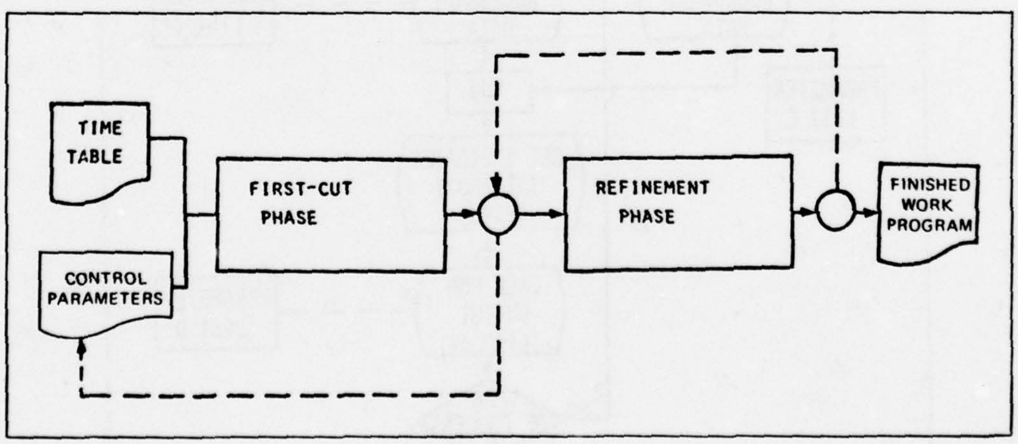


perienced some very embarrassing breakdowns, when the demand or the loading reached a saturation point. After lots of investigation people found out the reason for the breakdown was actually because an algorithm at a certain point lacked the human factor in it. The model simply does not have the insight to handle problems that a human scheduler intuitively can solve. So, an interactive mode was developed for this particular system, in which the human then used a graphics terminal, a microprocessor type, and could interrupt the functioning of the algorithm. He can get into the loop at two levels. One level is the parametric level. At this level the system is primarily steer-driven by the model; however, the human scheduler can supply parameters. For example, he can change the coefficients in the objective function, or even change the objective function and replace it by another one, or replace the statistical function or parameters built into the model. He can do it that way, but if that doesn't work then he can go into the heuristic loop — we call it heuristic — and actually make a one-on-one type of assignment.

The problem there, is can you switch back and forth? If the system is driven by some kind of optimization model and if you go in and change some parameter values arbitrarily, can you still jump back into the automatic algorithm — problems like that.

Another way of interaction is, again, from actual experience with a very large simulation system for assignment of vehicle operators. This happened to be in a company which is the largest transit company in the country involving thousands of operations a day. This was a system with automatic simulation and linear programming combined. It would take two hours to run. The result achieved something like 80% to 90% efficiency compared with a human scheduler. This would take two hours, and a human scheduler would take anywhere between three days to two weeks to finish a similar program. But the scheduler can achieve 100% efficiency. So what's wrong with that? Well, the problem was that when the computer generated a solution the human scheduler could not recognize it because he was not involved in the development of the schedule. Therefore he considered this as a black box approach and had no way to refine that. When the schedule was produced by linear programming, he couldn't trace the sequence leading towards the solution of the problem, so he simply would not use that. So, the man-computer approach then would divide the thing into two phases (Slide 26-3). The first phase we call the first cut. There we were able to sort out about 15 to 20 very critical parameters that would actually govern the solution of the scheduling problem, and these parameters, interestingly enough, happen to be the hardest for the computer to estimate. But, the hardest one for the computer to estimate, or the model to estimate, also happened to be the easiest for human judgment because of people's insight and experience with this sytem. So, by supplying these parameters, the system can consistently come out with a first cut solution, which is about 90% to 95% efficient, very quickly because it's more or less deterministic, involving no large mathematical programming. When you get to this point then the operator goes into a second phase, which is a refining phase. In this phase, the computer merely served as a bookkeeping machine to supply him with measures of effectiveness to tell him how well he's doing. He can do this arbitrarily based on his intuition and judgment until he is satisfied with the final product.

FIGURE 26-3 MAN-COMPUTER INTERACTIVE SCHEDULING PROCESS

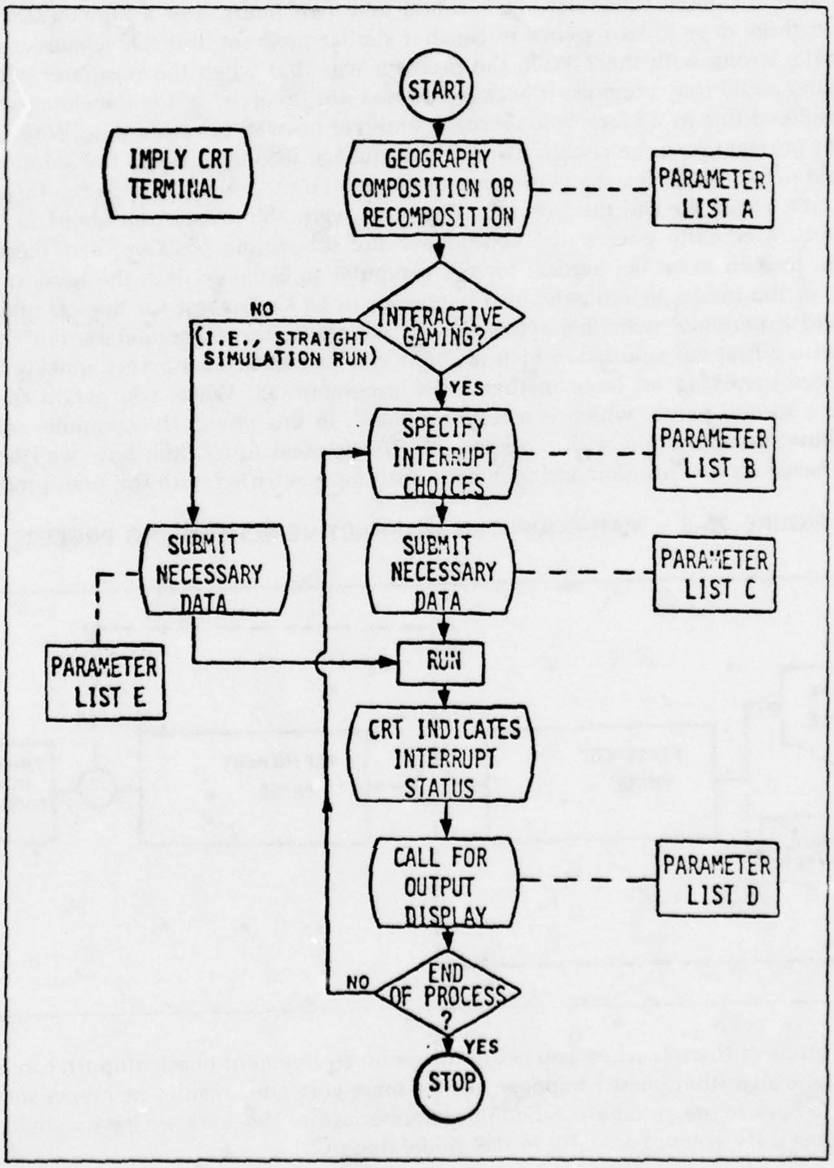


Of course, the problem there is when you get to this point (refinement phase output), can he go back to the "first cut" or automatic algorithm phase? Suppose he gets some very good results, he makes some changes, then he says, "I want to go back to the automatic schedule generator again." So, here we have a slight problem. Not so much in the computer software, but actually in the model design.

Let me quickly progress into the work we are doing for BALFRAM. Today, because of the time limitation we're not going to BALFRAM. I presume many of you know that model much better than I do. We have the father of BALFRAM here, Larry Low. Actually, I think he's the grandfather of BALFRAM. CINCPAC is the current user and I guess Colonel Bob Doty is here, so he can give you the view from a user's point of view.

Here, of course, we are only at stage one, the whole thing involving several stages, but here we are at stage one of a whole process where we envision that some of the process can be done at a microprocessor type of front end with a colored CRT, and that a decision maker or his agent, or analyst can use it, for his own part. At the beginning (Slide 26-4), the user will just compose the map at the CRT, and also provide certain static input to the system. At that time the person has a number of options to consider, and he should decide whether he wants to go for a very straight simulation run or wants interactive gaming. If he wants interactive gaming, then he has to decide what kind of interrupt he wants. He can state it very simply — every eight hours of simulated time I want the model to stop and I want to review the progress. Or, he can specify a condition: when the Blue ground unit, number X, has reached the attrition of 40%, or he can say when the FEBA has reached a certain place. He can also simply say when the computer simulation comes to a certain point, "program requires additional data." And if there's no data then the model will stop waiting for the man to supply the data.

SLIDE 26-4 ILLUSTRATION OF A MAN/COMPUTER INTERACTIVE PROCESS



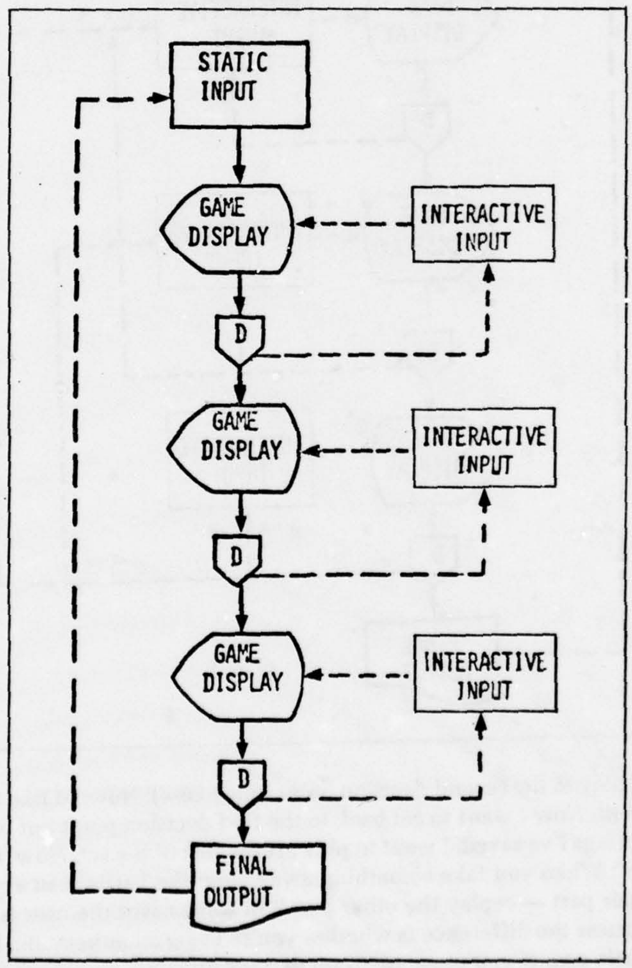
So, he will specify which of the above choices he wishes to make and supply the necessary data to the first stop and run the game. When the stop point is reached he can call for some output, which has several levels of aggregation and detail. After he examines the output he then might decide to make some other decisions. He can go back and change his choices and then go on to perform this process repetitively until either he decided he didn't want to wait until the final outcome because he has already received his information, or he can go to the very end of that particular experiment.

Now, I've given you some very crude examples. Of the problems we have encountered, and these are not actually in any particular order, nor are they all inclusive, the first one is the level of human control. When we say we put this human in the loop, how much control really do we give to the person? We talk about two levels so far (Slide 26-5). One is the macro level, the other one is the micro level. At the macro level, the human can change the function or the parameter values in the model, but he cannot really touch the detail values of the play. In the micro level then he can change anything he wants, but there are limitations. For example, he cannot suddenly in the middle of the battle change the order of battle or arbitrarily take a node out of the battle or put a new node into it. But, we really can't say he can never do that. He can put a new node in it as long as that node has never been interrogated. So, this again is not really a software problem. This really is related to the second problem; human judgment as opposed to automatic judgment.

Slide 26-5 — LEVELS OF CONTROL

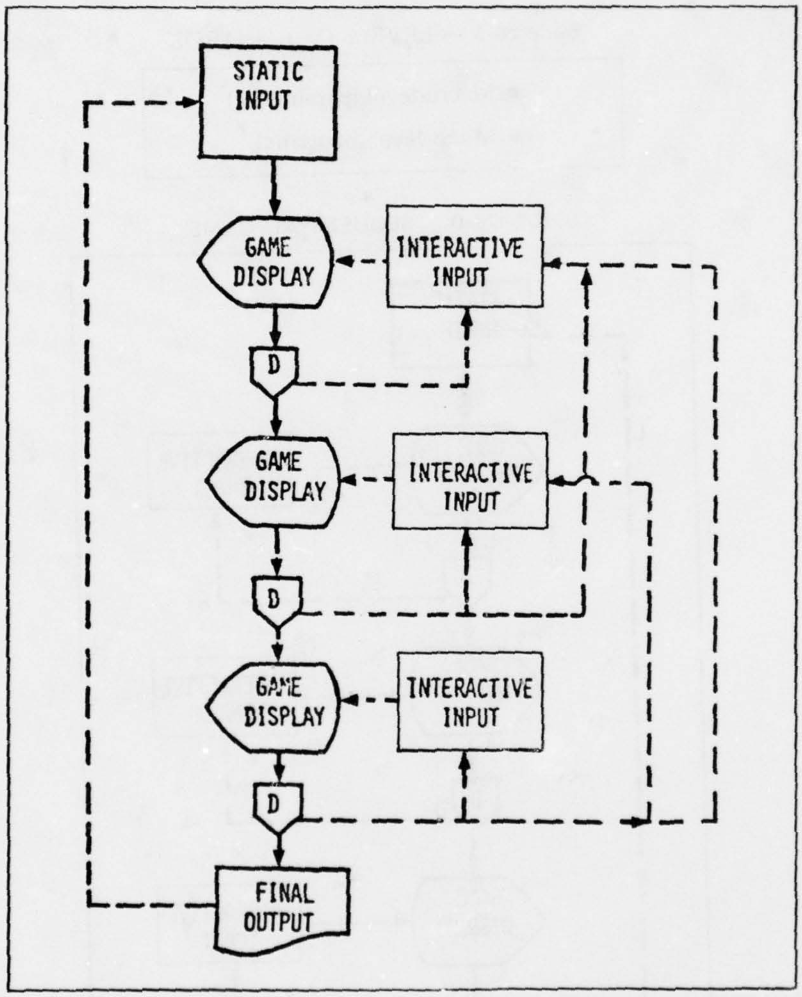
- Macro-level (parametric)
- Micro-level (heuristic)

SLIDE 26-6 SEQUENTIAL GAME



To give you some examples of how interactive techniques can be used in model exercises to explore alternative decisions (Slide 26-6) I think this is somewhat related to Randy's (Simpson) comment a few minutes ago. We are playing a game and you can decompose the model very simply by time segments. So, that's very natural in the simulation. You say, "okay, I can play forward to a certain point and I'll examine the output and I'll determine whether I want to replay that part." If you want to, you can go back and replay it to the same point again, and then you progress on to the next segment. So, you can very easily decompose the problem into time segments. Naturally, you can also do it other ways; you can do it geographically, hierarchically by levels of aggregation or hierarchically by the structure of decision making. W. nave such a system, but then when people use the system they get very bold and want to go back more than one step (Slide 26-7). In other words, I don't want to just go back to the immediate preceding decision point, I want to go back two or three decision points, or all the way back to the front and replay the whole thing. So, that's very straightforward. It's just a computer problem how to keep track of the history of the plays and the display of the decision tree and all the indices, but we found out the users will not stop there. They want to have their cake and eat it, too.

SLIDE 26-7 RECYCLING GAME

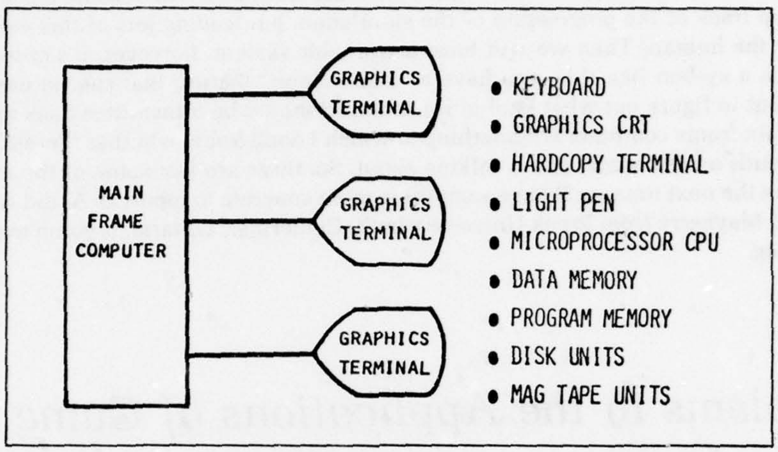


For example, I'd like to play to the second decision point (Slide 26-7). Now I'd like to save some of the results at this point and so I flag them. Now I want to get back to the first decision point but when I play it again I don't want to touch all the other things I've saved. I want to play off the rest of the set. Now, that's very hard to do. Can you still reach the optimum? When you take something away from the battle, you say, "okay, I like the output here, but I don't like the other part — replay the other part." In some cases the user is right, you can do that. In some other cases he's not. I guess the difference is whether you're using a mathematical programming type of optimization or a purely heuristic one. If you use something like the dynamic programming as presented yesterday

this will be very hard. When you say, "okay, the second time when I replay or recycle, I'll only use the remaining set, saving the other one," you will get a suboptimal solution.

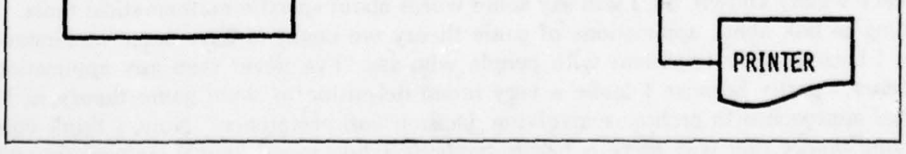
If you can successfully then decompose a model or game into various subgraphs or subnetworks, then it is possible to play the game in some manner like this: a simple chart that shows you have the main frame computer here (Slide 26-8). Then suppose this game can be played by three people, very much like what Al presented. You will have three graphics terminals, then maybe one person can represent Blue, the other one represents Red, the third one is the referee. Or you might have one represent the air force, one the ground force, and the third the naval force, and so forth. It's easy to have a setup like that, but it's not very easy to decompose a problem so that you can go into parallel operation. The only time you can have parallel for sure is when you haven't reached a FEBA, or the forces are moving but they haven't encountered each other. But that's not very interesting because you don't need interactive play at that moment. But once you reach the FEBA and you have interaction, it is not very clear how you can divide this thing up.

SLIDE 26-8 INTERACTIVE GRAPHICS TERMINALS



We encountered a problem, not only in the military, but in the commercial sector back in the 1960s when there was a trend to eliminate all the little computers from the subordinate organizations and concentrate all the data processing power in a centralized location. Especially in a theater command, you'll find the computer center has a very large machine like the WWMCCS system that runs on time share and on-line. So now we are talking about using, say, a front end computer (Slide 26-9). The advantage of doing this is that we might be able to use the centralized computer with a minimum amount of interruption. But if we say, "okay, we have a man/machine interactive system using the large computer, say a CDC 6700 or a Honeywell computer," I suspect that the first time you demonstrate this to the general, putting a terminal on his table, in a few seconds the picture will be lost, or you might experience waiting several minutes to infinite time for a return. This can be very discouraging. The

SLIDE 26-9 A DISTRIBUTED SIMULATION PROCESS



centralized system is not really set up to support that kind of interaction, but fortunately with the advent of microprocessors, (we have done an extensive survey of the microprocessor market) it may be possible very soon to use such a system both economically and technologically. What we plan to do then is to use the intelligent terminal "stand alone" feature to do lots of the front-end work. For example, you can compose and recompose a map, using a color CRT with a very low resolution of, say, 256 by 256. Really, it's a pseudographics terminal, as opposed to 1,024 by 1,024, but you can display a schematic map in 256 by 256 with color. This will solve some of the human factor problem, and you can do that on the front end, plus some bookkeeping processes. With the microprocessor today, you can get 64k bytes of memory as standard. Then you will only transmit the map to the central computer not in its picture form but in a matrix form, say a zero one matrix, which will maintain the isomorphic relationship with the original map and the main computer can do the processing in the batch mode. The user then will not even be aware that he's using a larger computer. He can interact solely with this front end. In the event that communication is broken we envision that maybe by this "stand alone" feature we can do very crude gaming. The problem then is how do you divide up the jobs? We have some ideas; for example, the larger algorithm work will be done in the main frame computer but whenever the algorithm can be replaced by human judgment, that can be done in the front-end computer. So, the front end will become a computer only to do the bookkeeping, to keep track of the progression of the simulation, but leaving lots of this very complex judgment work to be done by the human. Then we will have a workable system; however, it's crude.

I would imagine a system like this: you have a "stand alone" feature that can be used very efficiently to divide up the jobs. But to figure out what kind of information should be transmitted back and forth between the front end and the main frame computer is something to which I don't know whether there is a general solution. It really probably depends on the model one is talking about. So, these are just some of the interesting problems I throw out, and I hope the next time we'll have something more concrete to report as Al did with the TRW system.

Prof. Lucas: Jim Mayberry from Brock University in St. Catherines, Ontario, is going to speak on some of the problems of modeling.

27 — Problems in the Applications of Game Theory at the Theater-Level

PROF. JOHN MAYBERRY
Brock University

Some elegant insights into nonzero sum games and games of incomplete information

Everybody else has been talking about problems with game theory and with war gaming, so I'm not going to be unique in talking about the problems, but I'd like to set the stage a little bit more.

First of all, I'm glad to be here. Brock University is not a major center in the U.S. Government's war gaming activity, and consequently some of you may wonder why I'm here. There was a period yesterday morning when I was wondering that, too, then after Martin Shubik's talk I began to feel perhaps there was some point in it.

I was glad that Martin and Mr. Dunnigan were here to cast useful light on what people do when they're actually trying to do a war-gaming job for a customer.

I think that now there are lots of tools available, some of which are being widely used, and some of which are not very widely known. So, I will say some words about specific mathematical tools.

If we're going to talk about applications of game theory we ought to have some understanding of what game theory is; I become very impatient with people who say "I've never seen any application or any real use of game theory," partly because I make a very broad definition of what game theory is. I define game theory as "formal approaches to problems involving decision and preference." Now, I think you'll agree that if you define game theory that way there is hardly anybody who can say that it isn't useful to him. The ex-

ceptions arise, even though everybody faces problems of decision and preference, because we don't have to face them *formally.* I use the word "formally" rather than "mathematically," although I regard the words as synonymous; if I say I'm talking about mathematical approaches, somebody's going to ask "where's the mathematics? There's no epsilons and deltas, there's none of what Colonel Lemay used to call piccolo music, and if you don't have any of those funny symbols how can it be mathematics?" Somebody else is very apt to say, "I know that I don't understand mathematics, and no matter how you try you can't convince me; if you show me something that I understand, that proves it's not mathematics."

Well, sidestep that rather fruitless debate, and say "formal methods," rather than "mathematical methods," and then I think most of us can agree.

These formal methods, the problems of decision and preference, clearly include a great deal. All of pure mathematics, for example, is included in the trivial case of the zero-person game where there are preferences but no decisions; it's just a matter of finding out the good theorems. My view of game theory also includes all of economics and sociology and politics, so I cast a rather broad net when I talk about game theory. Maybe it's a little bit reminiscent of the "New Yorker's map of the United States," where Ebbett's Field is shown as about two and a half times the size of Texas. It may be biased, but that is how the New Yorker may view the United States, and that is how I view the world of game theory.

War games, game theory, any kind of models that we may look at, are fiction. Now, I have a suspicion that most of you are saying, "so what else is now," and a few are saying, "like hell it is!" I think that war games and analyses are fictional in the very important sense that they ought to be judged by the standards by which we judge fiction. We ought not to forget that they're fiction. Somebody was mentioning at dinner last night that when the first Tarzan films came out there were a few kids that broke their legs or broke their arms by leaping off barns and trying to do the things that Johnny Weissmuller had done in the films. It's very important not to forget that these things are fictional and, furthermore, not to forget that no matter how much detail you put in and no matter how much you work at improving the data base, they're *still* going to be fictional. You may make it *good* fiction; you may make it *convincing* fiction, which is not the same thing; but, you won't make it *nonfiction*.

A possible viewpoint is that the purpose of fiction is for training. Now, I recognize that this is a little bit philosophical, but I think it's fairly important. No one, I think, could ever say "I've done a better job this morning because of having read King Lear." At the same time I think it's perfectly possible to contend that a person who's read King Lear may understand what it tells you about the nobility and meanness which can coexist in a man's character and how that can destroy him, and that such a person may be wiser — better-trained for the difficult decisions and experiences of life — than one who has *not* read "King Lear."

It's difficult sometimes to point your finger at exactly where you use specific things you learn. I don't think that necessarily proves that you didn't use them.

A parallel here perhaps is the fact that the old masters, the artists whom we think of as really knowing how to paint the human figure, were also students of anatomy. Very few of them showed pictures in which the bones of the person they were painting were visible, and yet, they were able to paint the flesh and the muscles more convincingly because they knew where the bones were. I think this is the same kind of reason for which one ought to know something about the mathematical underpinning, the bones if you like, of what you're doing. It can make the result more convincing and more believable and more real. It can, as Poohbah said, "add artistic verisimilitude to an otherwise bald and unconvincing narrative."

Well, I think that perhaps one of the reasons for the success of the work that Mr. Dobieski was just talking about a few minutes ago is that that work was done for training and it was quite clear soon after the beginning, who was going to be trained and to do what; for that reason they had a task which was *possible* to do very well. I'm not attempting to diminish what looks to be a splendid accomplishment, but I think that many of us have been in a gaming situation, where the original problem was asked in such a way that a splendid result would simply have been impossible. Partly, perhaps, because it wasn't clear *who* was going to be trained to do *what*.

Of course, if you're not clear in what you want to do, it's never clear how to do it, or, as the White Queen said, "If you don't much care where you want to get to, then it doesn't matter very much which way you go." Abraham Lincoln said "if we but knew where we were, and whither we were tending, we could more easily discern what steps to take."

Now, having, I hope, got you all convinced that there might be some point in looking at what mathematicians have been doing off in their ivory towers, then I think I will try to follow up on Martin Shubik's contention that they have indeed been doing something. So, what I want to do is to use this classical educational device (blackboard) to show a little bit about what people can do when they look at games.

First of all, what you can't do: you can't do anything useful with the zero-sum two-person game. I feel this very strongly. I shared in the excitement and the enthusiasm a lot of people had a few years ago that

game theory was going to revolutionize the world. I'm still convinced that it's going to, but now I perhaps express that a little more definitely in the future tense than I did then.

A zero-sum two-player game (I'd rather say two-player than two-person for the obvious reason that very often the players are not persons, but countries or nations or groups of people) implies that whatever one person loses — (see how seductive it is!) whatever one player loses, the other one will gain. In fact, if you look carefully at the definitions, the losses are not to be measured in objective quantitative or physical terms but are to be measured in psychological terms according to the preferences of the players. So, it is a basic assumption of the two-player zero-sum game that whenever one player feels that he has lost something the other player will feel that he has gained an exactly corresponding amount. Now, on the face of it that's nonsense because these preferences in the one case are inside the head of player number one, in the other case they're inside the head of player number two. Apart from accidental coincidences, I can only see a situation where a zero sum game would be applicable in cases where there is some objective physical thing that gets transferred, such as, for example, money — when people play games for money. Then, *if* the house doesn't take a cut and *if* you don't take the price of the evening's beer and pretzels out of the pot, *then* the amount of money which one loses is exactly the same as what the other one wins. Even then, that doesn't necessarily mean that the preferences are equal, because as you all know, nonlinear preferences for money are quite typical.

If anyone doubts this I will give a completely rigorous proof. Insurance companies exist. Obviously, the mathematical expectation of a man who buys an insurance policy, is negative; if it were not so the insurance company couldn't pay for their computers, their clerks, and, of course, the profits to the widows and orphans who own the stock. Does that mean it's irrational to buy insurance? No, because if I lose my house or my car that's a loss which I would have a very hard time restoring. My utility for a loss of a large amount of money is much more negative than the amount of money itself would suggest.

In fact, I have a nonlinear utility for money, certainly when the amounts get large. This implies that the only case in which it's reasonable to look at the money which is transferred as a proxy for utility is when the amount of money is relatively small.

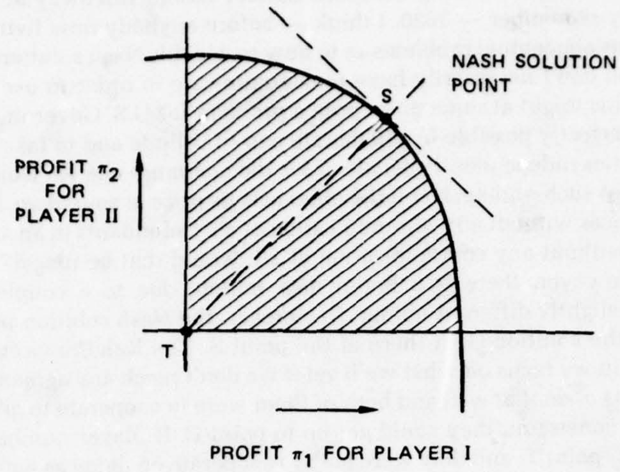
Okay, so somebody might say then the two-player zero-sum game is only applicable to cases where two people are playing a game for money, for stakes they can afford. But no, not even then, because when you're playing for money for stakes you can afford, generally the pleasure of the game has something to do with it, too, and consequently that's not zero-sum either. We all know that there are some people with whom we would not play a social game of cribbage or bridge, simply because we don't like them.

So, there is no such thing in real life as a situation which is well modeled as a zero-sum two-player game. This is not to denegrate the value of the zero-sum concept without which the more sophisticated notions cannot be formulated. Much of the criticism directed against attempts to use game theory practically may be due to uncritical efforts that employ zero-sum games where they are not appropriate. There are as it turns out, two very important uses of the zero-sum two-player game. The first is in a linear programming situation, when there's actually only one player; the other player doesn't exist, or isn't important, or isn't regarded as having much impact on the game. For example, in the case just mentioned of optimizing a transportation network by linear programming, there was no one who particularly wanted the costs of the transportation network to be maximized. Nevertheless, the mathematics were the same; the linear programming problem which was solved did correspond to a competitive game.

The second way a zero-sum game can be used is as a mathematical foundation, to help you get a handle on something which is not zero-sum. Now, we can draw a picture, as in an elementary economics course, and we call our players I and II (Slide 27-1). There will be some set which I'll draw for the moment with a nice smooth curve, and this is the feasible region, the set of possible outcomes which could be attained by these two players. The boundary curve defining the feasible region reflects the fact that total quantity cannot exceed 100%. The nonlinearity of the boundary is caused by the non-linearity in utility functions of both players. If we had a "default" or "no-agreement" point, we could regard that as a kind of zero from which we might negotiate. One of the characteristics of *economic* competition and cooperation, as distinct from *military* competition and cooperation, is that in economic situations there is, generally speaking, a default outcome. If we go into a market hoping to buy some bagels, presumably somebody came in there hoping to sell some. (Otherwise, of course, nothing happens.) There is a perfectly well-defined no-agreement point, namely, the seller sadly takes his bagels home with him and I sadly go home without any bagels and take my money home with me. So, let's for the minute look at an economic situation where there is a naturally-defined no-agreement point. In a paper about cooperative games, published in 1953, (and perhaps not as widely known as it should be) John Nash showed that, under certain assumptions, there is a certain point which would be a rational outcome under cooperation. It happens to be the point where the product of π_1, the profit of player I, and π_2, the profit of player II, is maximum, and that turns out to be a unique point under the assumption that the curve is convex. (The convexity means that the players are

willing to play mixed strategies if necessary; there are other philosophical considerations here which we'll postpone.)

SLIDE 27-1 NASH SOLUTION POINT



Nash says there's a perfectly definite cooperative point, once we've agreed on what will happen if we agree to differ. In other words, we come to the negotiating table, and we try to decide what to do; if we can decide on something, fine, we do it (providing, of course, it's within the feasible region). If we can't decide on something then we leave the negotiating table and what happens is this no-agreement point T. Because there is something above and to the right of (Slide 27-1) the point T, there is a possibility that both of us may profit; presumably that's why we came to the negotiating table. So as far as the qualitative discussions are concerned the only thing we have to know is that the solution point S is uniquely determined, and that its location depends only on the local properties of this curve, it depends only on the slope of the curve at the point S, which has to be the negative of the slope of the line ST. It turns out this is a reasonable thing to do and has nice mathematical properties. It also has some intuitive appeal. When described to business administration students as a reasonable way to make negotiations, they often find it entirely acceptable; in fact, some of them find it extremely exciting. I once had a student say that he wished he'd known about the Nash solution the previous summer, when his father was trying to decide whether or not to buy a drive-in movie; that, I felt, was one of the greatest compliments that any of my students had ever paid me. He really would have used the Nash solution! Now, of course, whether his *father* would have used it is another question!

However, there are still several questions about this Nash solution, even if we're convinced that, in principle, this is the way to negotiate.

One of the questions is, "when does this negotiation take place?" Within the context of the theory of bargaining games and non-zero sum games, that has a very clear answer, because negotiation takes place (if it actually takes place at all) *before* the game starts. Most game theory deals with finite games which have a beginning, an end, only a finite number of alternatives, and so on. It turns out the assumption that the game has an end, isn't so serious. There was some discussion of that yesterday. But, really, the ending of the game isn't such a serious problem for the analyst. Of course, it's a very serious problem for people who are having British troops in Northern Ireland, or American troops in Vietnam, or are doing something else which they'd just as soon stop thank you very much, if they could only figure out a way of stopping. But, mathematically, if one can discount the future in a suitable way, the infinite future of the game often doesn't pose a very serious mathematical problem. I see Randy (Simpson) looking very shocked.

Mr. Simpson: Practical problems, too, but go ahead.

Prof. Mayberry: I'm talking now about developing mathematical tools, and I'm not here to pretend that they can solve all problems, but *not knowing* about the mathematical tools can *create* problems, and that's what I'm hoping to help us all avoid.

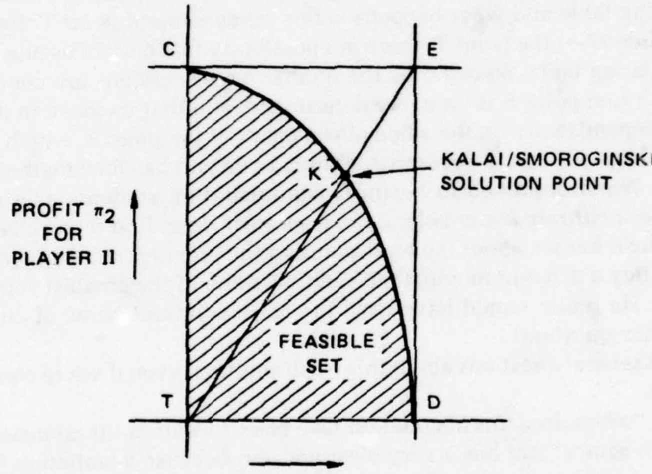
The more serious problem is caused by a game which doesn't have a beginning. As far as I know, nobody knows how to analyze a thing like that because it violates the whole notion of the strategic form, or the normal form of the game, where you simply list all the strategies and then you imagine yourself choosing among the strategies before the game begins. If the game doesn't have any "before," you can't do that, and, of course, that was perhaps the most troublesome thing about Vietnam where we became embroiled in spite of saying "never

again." Even after that had happened before their eyes, the British become embroiled in Northern Ireland. You could make a fair case for saying that the Vietnam war started some thousand years ago when the Indo-Chinese began to rebel against whoever it was that happened to be sitting on top at that time. You could make an even better case, I think, for the fact that the war in Northern Ireland was in full sway at the time of the battle of the Boyne, some historian may remember — 1620, I think — before anybody now living could have done anything about it. So, we have serious conceptual problems as to how to use this Nash solution. But nevertheless there may be some value because you don't necessarily have to communicate in order to use the Nash solution.

As a side issue here, this might at some point pose a problem for U.S. Government attempts to enforce anticartel laws, because it is perfectly possible to collaborate and to collude and to take cooperative actions without communicating; if two parties independently decide, if we did communicate we would reach this solution where one company sets its price at such-and-such and the other sets its price at so-and-so, it is entirely conceivable that they might choose those prices without any communication at all. Defendants in an antitrust suit might claim that it was possible to do this without any communication at all; would that be illegal?

Okay, it turns out, however, there is a brand new notion, due to a couple of men named Kalai and Smoroginski. They have a slightly different notion. Remember, the Nash solution point is determined primarily by the local properties of the solution right there at the point S. The Kalai/Smoroginski solution is based on a different construction. Again we focus on what we'll get if we don't reach any agreement — point T in Slide 27-2. Now, if player number I was to do that well and both of them were to cooperate to give player number II as much as possible, subject to that constraint, they could get up to point C. If player number II was to do as well as he could at the "no-agreement" point T, and they were both to cooperate on doing as well as possible for player number I, they could attain point D.

SLIDE 27-2 KALAI/SMOROGINSKI SOLUTION POINT



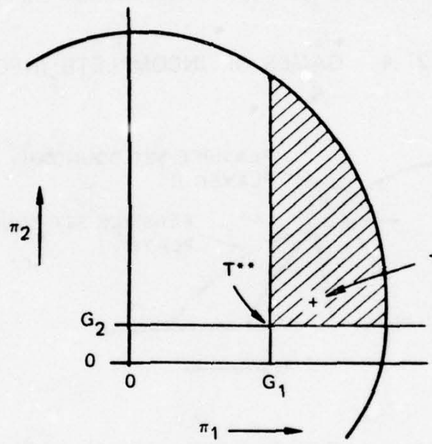
Again, you would expect both of these numbers to be positive since the reason why we're talking about negotiation is because of the possibility that both of them would improve, so then, this rectangle CTDE is filled out, the line TE is drawn, and this point K, where the line TE meets the boundary of the feasible set, is the cooperative point. It turns out that this solution also has nice mathematical properties but (of course) different ones. Nash took four axioms, which I and many other people found entirely convincing for twenty-plus years, and produced his notion. Kalai and Smoroginski took three of those axioms and one different one and got a somewhat different solution.

It turns out that in a lot of cases these solutions agree, but in some cases they don't, so it really is a different thing. My own feeling is that, for anything military, I can accept the idea of the Nash solution somewhat better in those cases where they differ. To explain my reasoning, let us suppose that, before a war begins, both parties have a pretty clear picture as to how the war would end. (Of course, it's generally not the case; the reason why the war starts is probably because both sides think they're going to come out on top.) But, suppose they have a pretty clear picture how the war would come out if they fought it. You may have heard war protestors saying "if all the money that was being spent by the United States on the war were spent on something else — whatever you happen to be in favor of at the moment, possibly even on foreign aid to Vietnam — look how much better everything

would be and the United States would be no worse off." Well, it's that kind of a question — "what would happen if we took everything that we were going to lose in the war and instead spent it in helping our potential opponent?" Well, it's not only difficult to answer that question, it's even pretty damned undiplomatic to ask it. But that's the question that one has to ask in order to get the Kalai/Smoroginski solution, and, therefore, my feeling is that in a military situation it probably is not as useful as the Nash solution.

Well, there are a number of other questions which come up in relation to this kind of essentially negotiated solution. One of them is "What is that threat point (the no-agreement point)?" In a military situation it isn't obvious what is the threat point, there is no such thing as a uniquely defined result of disagreement in a military situation; it depends not only on who wins, but on when the loser gives up or when the winner decides to stop pounding on the loser. So that, what you're faced with conceptually at least, is a situation where we don't know what will happen if we don't reach agreement. Well, one way of determining a possible threat point, as mentioned yesterday, is by this procedure. First, the players focus on player number I's payoff, and II temporarily adopts the goal of hurting player I as much as possible. We have created a zero-sum game simply by fiat, by both players paying no attention to what II actually wants. There is a solution to the zero-sum game, which says that player I can always guarantee that he gets at least so much (shown as G_1, on Slide 27-3) and player II can, if he chooses, guarantee that player I gets no more than that. Second, they can focus on the goals of player II; player II can simply play to defend himself and player I can play to hurt II as much as possible. Then we would get a horizontal level here; player II could ensure that he would get at least that much (shown as G_2 on Slide 27-3) and player I could ensure that II would get no more.

SLIDE 27-3 RAIFFA AND SHAPLEY "NO AGREEMENT" POINTS



Suppose now that both players play defensively; then, player I, by playing his defensive strategy, ensures that he's to the right of G_1; player II, by playing *his* defensive strategy, ensures that he's above G_2; so we know we're going to end up someplace in the shaded region of Slide 27-3 — and that point T*, wherever it is, might reasonably be taken as a "no-agreement" point. *Point T* is, in fact, the "no agreement" point proposed by Raiffa.*

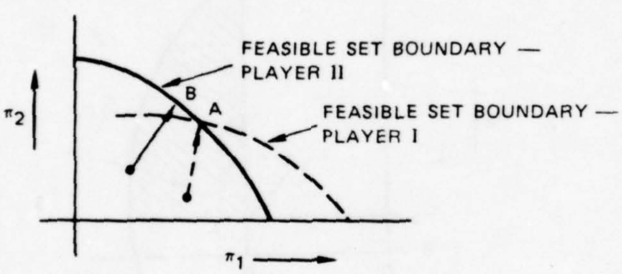
A simpler notion, due to L.S. Shapley, is that the point T** — where I gets G_1 and II gets G_2 — be used as a "threat point," or "no-agreement" point. Nash himself proposed the use of a still different construction, which has for me actually greater appeal but is too complicated to describe now. Details for this can be found in Nash's original paper in *Econometrica* (1951). Whichever of these three notions we accept, we get a definite point to work from, after which one can imagine the kind of negotiation described above.

Well, I think one can fairly easily imagine a commercial situation in which the facts were well enough known and well enough agreed between the two players that this kind of analysis could be done, even possibly could be done by each player separately and they could decide what they expected to happen if they negotiated face to face or if they simply decided to behave in a sort of cooperative way. It's not nearly as easy to imagine this being done in a military situation, because, of course, the crucial question, that people generally have no answer to at all, or only a very sadly inaccurate answer when they begin a conflict, is "how is all this going to end?" Very often they don't ask the question, but if they do ask it, the two antagonists are likely to give very different answers.

I have the impression, for example, that when the Japanese attacked Pearl Harbor, it was their plan to attack Pearl Harbor, destroy the fleet, then sue for peace with the United States; they hoped to make some fairly favorable bargaining terms which would, for example, amount to leaving the greater East Asia co-prosperity sphere under the influence of Japan, and that was the plan which the Emperor had agreed to. But then after the attack was even more successful than they had imagined, the general staff decided, "gee, this is fun, let's not quit while we're ahead." So, the idea of foreseeing how something may end up, what would be the result of a conflict, is something which is not well done at all. It can't be well done I think, and therefore we really need to look at the possibility of games with incomplete information.

Now, "games of incomplete information" is a subject which covers a number of very exciting pieces of mathematical territory. I think they are very likely to give rise to things which are going to be practically applicable in the future. Right now it enables one to think about non-zero-sum situations, and about competitiveness in military situations, but not very quantitatively for two reasons; one is that there are many different models and the other is that almost all those models share one regrettable property — the computational and mathematical difficulties of finding the answer, even to very very simple problems, are horrendous. That is not going to be cured by bigger and faster computers; it's not that kind of problem, it's more conceptual than computational. But, nevertheless, it's possible to say what would happen if player I thinks that that would result if they didn't agree (Slide 27-4), and that that is the feasible set (dotted curve of Slide 27-4) and if player II thinks that is what would result if they didn't agree and he thinks that this is the feasible set (solid curve). Obviously, they can't do anything without a certain amount of communication, but the type of negotiation which they might agree on is obviously very different. Instead of saying we will do something which will result in a particular definite result, they may want to negotiate on what they will do, or they may want to negotiate on how good it will be for them. In other words, one person might go to the negotiating table and say I want you to take that action, or another person might go and say I want you to do something which will be at least this favorable for me.

SLIDE 27-4 GAMES OF INCOMPLETE INFORMATION



So, the result which they might agree on would then be something which might be evaluated as being as good as Point A (Slide 27-4) by one of the players, and the other player might very well see that the profit in the discussion had been Point B.

There is one other possibility. This type of game of incomplete information permits negotiation as the prototype of all bargaining situations, and for some reason whenever I think of bargaining I think of one person wanting to buy, and one wanting to sell, an oriental rug. We can formalize that situation very nicely; we must look at how much each player can believe about what the other one says. One of the interesting formal results that came out of some work that was done by the Arms Control and Disarmament Agency a number of years ago, and I'll just leave this with you as a provocative last word, was that in a situation where lying is possible then a person may very well find that he cannot communicate information even if he wants to. (He can say it, but if it was possible for him to lie, he may not be believed). On the other hand, in a situation where lying is *not* possible he may very well give away information, even under circumstances where he doesn't want to. For an example of the first case, say I want somebody to believe that this really is a gold coin and I'm starving to death and I need the money and I really want you to buy it and it really is a gold coin, but he knows very well you could say that even if it wasn't, consequently he won't believe you and so you may end up starving to death. Or, alternatively, if it's not possible to lie, if you have a lie detector or you're under oath or something, you may very well have a fake gold coin and you may want to say "this is real gold, buy it;" if you can't lie you won't say that, and when he asks, "Is it real gold?" if you don't have the freedom to lie you can't answer and consequently he'll know that it's not. So, this I feel is an extremely important insight and it also has some very interesting applications if you think about the way the diplomats communicate information. It's not through the official notes, it's through people talking at

cocktail parties, and the ambassadors talking to somebody on the other side and then giving information informally. But this question about how easy it is to lie, and how easy it is to tell people what you want to and to avoid telling them what you don't want to, obviously becomes very important.

There are lots of other things going on in the area of games of incomplete information, which are obviously of tremendous importance to a military situation, and I'll simply mention the two main categories: first, repeated games where the players can infer what the game is, and what the other person thinks, or what he knows, or what he wants, by watching his behavior in repetitive situations. That has a very beautiful mathematical theory, associated with the names of Aumann, Maschler, Mertens, and Zamir.

Second, there are the games of bargaining with incomplete information, and I feel that there is real pay dirt there, but at the moment, the computations are such that it's very difficult to get answers to simple problems. So, I suggest that as far as the bargaining games are concerned, you keep your eye on journals like the *Journal of Conflict Resolution* and the *International Journal of Game Theory*.

Prof. Lucas: And "*Mathematics of OR*;" charity should begin at home, shouldn't it!

Prof. Mayberry: Yes, "*Mathematics of OR*" also, and take a look at things which refer to bargaining games because if it hasn't happened yet it is very likely to happen in the next few years, and it will be of great interest.

Prof. Lucas: There are two Kalais who published in game theory, so let me just add a first name, Ehud, to this one. He's currently at Northwestern University, although he seems to also hold an appointment at Tel Aviv, and I don't know by which one he's given tenure. He's Israeli and his wife is American.

Panel Discussion — Session IV

Prof. Lucas: Over the last seven years there has been quite a bit of development in games of incomplete information and in the repeated or stochastic games involving such concepts of averages. Here are two very current references — these are fairly mathematical, but they may give you a place to start and then work back through the literature of references and things. There were three papers that had a lot to do with getting the theory going, by John Harsanyi in the mid-1960s in *Management Sciences* (I'm not sure of the exact date on that). Some of this work goes back to some Arms Control Disarmament Agency reports. Some of the nice little examples in these reports I do not think have ever gotten published since. Some may have appeared in Tom Saaty's book. There are a couple of errors right out of the report and reproduced in the book. These reports are unclassified, but on the other hand, they're ones that you write and ask for and they don't send you. The latest that I heard is that Tom Saaty has a lot of these in his garage, so perhaps you could write him directly. He was the scientific officer of some of these projects at the time.

In our panel, and Martin Shubik take note, we have quite a few young people here. We'll have each panelist speak for a few minutes and I hope we still have some time for questions without getting too far behind.

Lola Goheen is at SRI International. She has her Ph.D. from Stanford University, and she has worked on some of the analytic problems, some of the closed form solutions for some repeated games and things.

Dr. Goheen: I just wanted to relate to you people today some of our experiences at SRI with solving multistage games. We wanted to optimally allocate the air resources in the BALFRAM simulation model, and, as you know, BALFRAM, depending on how you define your terms, handles naval, air, and surface engagements. The thought was that it would be nice if you could optimally allocate all the forces and to start with air since they seem most amenable to this kind of optimization.

We looked at various methodologies for solving multistage games. We looked at DYGAM, the Lulejian methodology — the air optimization part of their methodology of La Grange dynamic programming which is the methodology they used in TAC CONTENDER to optimally allocate air — and we looked at the closed form solution which Berkowitz and Dresher did in the early 1960s and late 1950s. So our task then was two-fold: to evaluate the decision games and pick one we thought we could interface with BALFRAM — to pick a good decision scheme, one that was sound and one that seemed feasible, and then to interface it with a large computer simulation.

I'm talking about the problems we had because that's pretty much the way it went — one problem after another. Some of the techniques appeared erroneous in that it was pretty easy to pick examples that Lulejian failed to solve and the La Grange dynamic programming failed to solve and the fact they not only didn't solve them but they failed to stop so that they would alternate back and forth between nonoptimal solutions. I did some of this work and others have done it, Karr for example.

Some techniques appeared infeasible to interface with BALFRAM, which was a slow running simula-

tion model at the time. It's being improved today. We could probably go back and look at DYGAM, for instance, and OPTSA. I didn't mention it, but we also looked at that — The OPTSA model from IDA, the early OPTSA.

DYGAM troubled us because the handling of mixed strategies isn't very clear, and it was also slow running. CINCPAC was using BALFRAM for planning purposes, and other places were using it, and they just didn't have computer time or facilities to support the computational requirements. The OPTSA model, the dynamic programming OPTSA also ran slowly at the time. It has since been improved and streamlined. So, we chose then, for better or for worse, to try to do a closed form solution, getting a very aggregated air/ground war model, and then just trying to solve it as Berkowitz and Dresher did, working backwards with closed form dynamic programming. To interface it with BALFRAM would have been fairly easy, because that would have just required table look-ups for the function of plugging in numbers once we had solved the closed form game.

So, what we had then was to determine our closed form model. The model that we wanted to solve with the N stage methodology, turned out to be one of our biggest problems in that the model that we chose was highly aggregated but yet was still much too complicated to solve analytically. The model that appeared desirable apparently couldn't be solved (by us, anyway) although we spent a lot of time at it, and when we did finally find a tractable model, it was so far from being realistic that it was under pretty heavy attack from even our own people, so it took out a lot of the incentive for struggling along with it.

Of course, generally, when you're trying to determine your model, you have to worry about your payoff, and we decided to use the difference between Red and Blue survivors for our payoff. We started with the state equation, with the transition equations from one state to another being the Lanchester Square Law, because it was fairly tractable analytically rather than being that realistic. So, at least, it was the Lanchester Law and BALFRAM could handle the Lanchester Square and Linear Laws.

One of the problems apparently in developing our model is that the interdependencies from one stage to another really interact in a very complex and conceptually incomprehensible way. When we finally get a model that appears tractable we find that we have difference equations. We find that instead of the Red and Blue survivors being given by the Lanchester Square Law, the payoff is ordnance delivered, or the number of aircraft surviving, which doesn't clearly relate to how your ground war is going.

What we wanted to do was to use an objective function that told us how the ground war was going. What we finally wound up having was something that showed us how many Red and Blue aircraft were left, and then we had to have such factors that could somehow tell us based on how many aircraft we had, how we were doing on the ground which, of course, is a discouraging turn of events. So, I just want to leave you with that bit of perplexity; the work we are doing is still in an unfinished state. We and the sponsors are just sort of sitting back and trying to figure out where to go. The reports are available from SRI and DDC, or if anyone is interested they can drop me a line or ask me about them.

Prof. Lucas: Our next panelist is Jeffrey Grotte. Dr. Grotte has degrees in applied mathematics from MIT and from Cornell University, and he is with IDA.

Dr. Grotte: I've sort of been wondering, as I've been listening to these game theory talks, what hat I should be wearing here. I'm not a theater-level modeler. I don't use them. I don't develop them. I work mostly in the strategic area where modeling is a lot less controversial, probably because there are fewer experts in it. But, I think I've seen a position to adopt — I'm going to be the game theorist, but a one-time game theorist. That hat is a little musty — who is on the inside, who's going to have to implement these things in studies, those are the questions. I have to take into account the structures of game theory, the practical concerns of the customer who is not going to be too happy to hear a lot of theory if I can't deliver what he wants, and the position of myself as the analyst, running up to the computer, solving the problems of the project leader who is getting worried and all that sort of thing.

Game theory, as you've heard, has a number of things to offer. It provides paradigms, models with which to analyze conflict situations, and from that point of view it's very good. There are many such games: as Professor Shubik mentioned, there are the Colonel Blotto games, N-person games, and Professor Lucas mentioned nonatomic games. Unlike Professor Mayberry, I think two-person zero-sum games do have some use, particularly when your other player is an unknown quantity. You don't know what his utilities are going to be and you can substitute a devil's advocate for that person. But, I see too few of these being employed, and, in fact, only one or two zero-sum types. Moreover the N stage game is being forced to do things for which I don't believe it is appropriate.

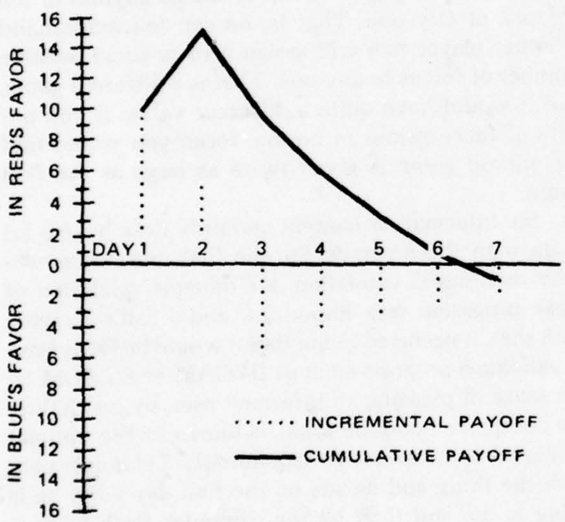
As we heard yesterday, the games are intimately connected with the information structure, and a good example, again, was the N stage game. One of the assumptions behind N stage games, or one of the assumptions that is consistent with N stage games, is that both players know when the war will end before the war starts. I don't think it is enough just to say, "Well, we'll discount the future because we have to solve all those games." We can't solve one anyway, as we heard that DYGAM can only solve six or eight

stages, and OPTSA only three. We can't run a game up to 100 days and see what happens.

This knowledge of the future that affects every step you take along the way — you make your decisions at stage three because you know it's going to be over at stage five — is, I think, inimical to warfare and is an area in which we need a great deal of work before we have confidence in the validity of our games.

In the question period last night, I think Jim Taylor was one, we heard some people say that the length of the game is going to strongly affect the behavior of the two players. And I would add that there's another thing that is going to affect the two players, and that is that they both assume that the game ends at the same time. If the players assume that it's not going to end at the same time their actions might be different. I have an example that I've been carrying around for two years waiting to show it to people, and you're going to see it (Slide C).

SLIDE C GAME OUTCOME EXAMPLE



Prof. Mayberry: Most examples of game theory get carried around a lot longer than that.

Dr. Grotte: I know, I feel privileged. This is a simple Dresher air war game. You're all familiar with it. Perfectly symmetric sides, Blue and Red. I solved it to seven strategies because this game can be solved explicitly. The measure of effectiveness is the difference of cumulative ordnance delivered. I'm sure a lot of you have seen this, and on this side is in Red's favor, and on this side is in Blue's favor, only Red has decided to play the strategy to correspond to a two-day game. Now, what happens when the game goes longer than that is that he just continues doing what he was doing the day before because he's very puzzled, but he wants to do something. Blue plays a seven-day strategy. Now, to be sure, Blue wins after day seven, but if you're Blue I don't think you're going to be too encouraged by that kind of result for your war. I think that's the case of your F-15s skimming along the ground. Red has pushed them all down, but the FEBA is in the Azores by day two, but by golly you can get it back there by day seven.

I throw that out just as something to think about. I think these are the sorts of things we have to worry about with N stage games, and I would hope that the people who use them at least give these sorts of things some thought. What I would like to see, and what you see more of in the differential games, are games where the terminal condition is not time, but some other condition — movement of the FEBA to here, a cumulative force delivered. I think you could also solve these pretty much the same way with dynamic programming techniques, but they're not as easy to do. But, I think that would be a great advance as far as generating studies that employ such techniques.

One of the other things that game theory has to offer are game theoretic solutions. The reason game theoretic solutions are valuable is because you can understand how they behave. When I see something like DYGAM being used to generate nonoptimal solutions, good as they may be, they're nonetheless nonoptimal, and I am very concerned because I do not know how those solutions behave. I would much rather, from my own standpoint and from the standpoint of communicating to a customer how a heuristic works to be able to say this is what it does here and this is what it does there instead of saying I have a nonoptimal solution that I got from a very sophisticated approximate routine that solved the war backwards. That's my own opinion and I think many of the customers might feel that way also. An additional value would be simple heuristics that would be transparent; you can explain them. You hear a lot about transparency, putting your assumptions up front.

ATACM, which we heard about, I think is in the right direction of more detail. The single notional airplane just won't fly anymore. You can't get people to accept that. They want to see all this detail, even in simple sorts of analyses.

I would like to point out in connection with ATACM that those max-min strategies and min-max strategies are not unique, so that min-max versus max-min value is not necessarily unique. It can be anywhere between the Red maximum minimum value and Blue's min-max value. So, all you are really getting out of that are those ranges, those bounds. They can be useful if they are close together. The result of this force structure gives you a range here and the other force structure gives you a range there. If they are far apart, I don't know what to say about them, but

I think you should be very careful when someone throws something out like that. Min-max versus max-min strategies as the result of conservative play are not unique.

As a final point, we heard that in connection with ATACM that mixed strategies are not good because you only want to play the game once. I think anyone that believes that should perhaps go back and see what mixed strategies are really for. They are for games that you only play once, and a good session with Von-Neumann and Morgenstern wouldn't hurt, I think.

Prof. Lucas: Our next member of our panel is James Falk. He is a Professor of Operations Research at George Washington University. He has an undergraduate degree in engineering from Detroit, and has a Ph.D. in mathematics from the University of Michigan.

Prof. Falk: There are two reasons why I am here today. One, the first person I ever had for a calculus course was Bill Lucas, in Detroit. Two, a few years ago Jerry Bracken at IDA asked me to take a look at the TAC CONTENDER model, which I did. I am still not sure whether I should thank him for that or not, but there certainly are a lot of interesting problems with this area.

I would like to mention some problems that I have noticed with sequential games. Jeff mentioned one. There are some others like TAC CONTENDER, for example, the original document for which I thought was very, very well done. The difficulty with TAC CONTENDER was the the problem TAC CONTENDER methodology set out to solve was quite well detailed but TAC CONTENDER didn't solve it. It solved a different game problem, one in which inventory left after a day's action was the expected value of the inventory coming from the previous day's action. Okay, let's leave that alone. Maybe we can live with that.

Another problem with TAC CONTENDER was that it could, as Lola Goheen has already mentioned, oscillate between nonoptimal solutions. That's not too nice. What is even worse is if it settled on a optimal solution and announced "here I am, and this is optimal," it was possible the solution was not optimal, and roughly this corresponds to the fact that if you are optimizing a nonconvex function you could settle on a local optimum, and say, "ah, that's it because it's better than anything nearby it."

Well, DYGAM does try to solve the correct game, and intuitively I feel that that's probably the correct approach if you want to solve something with any kind of detail. Of course, the difficulties are obvious in that it runs for a while, but John Tomlin is going to help out there, and it does produce approximate solutions and while they have done some things to estimate the error involved, there is still the uneasy feeling that it might not be real close to optimal.

ATACM, like one operational mode of DYGAM, solves the sequential max-min problem, and one way to think to about this max-min problem is the following: You let one side move, and the other side react to that move, knowing exactly what that move was. Well, this is tantamount to assuming that one side has perfect intelligence, that its spies are very good, and the other side was none. That's to get one value, say, the lower value. To get the upper value, you switch the roles. Okay, if you looked at the simultaneous game, you would, in effect, assume that neither side has any intelligence value so you get some sort of an intermediate value.

Well, Martin Shubik mentioned this when he was talking about the sensitivity of the final value to the information content. You could think of the Colonel Blotto game where, let's say, there are three cities and one side wants to defend each of three cities and the other side wants to attack. Well, imagine one game where both players have to make their decisions simultaneously and in complete ignorance of what the other side is doing. Okay, that's one game.

Imagine another game where, let's say, the attacker has an intelligence force which can predict whether or not player two will assign anyone to the defense of city one. That is, he can tell beforehand whether player two will assign zero or some positive number of forces to city one. That is a different game, and it would have quite a different value. If you put both of those games in normal form, you would find the second game is about twice as large as the first game.

So, information content certainly does have a lot to do with these games. On the first morning somebody mentioned validation. He thought validation of these programs was important, and I can't disagree with that. It occurred to me that it would be fairly easy to validate a program such as DYGAM or ATACM, in the sense of pleasing an informed user, by just letting the computer compute what it thinks of the optimal strategies and then move sequentially. Let a user play with the thing and decide on the first day what he is going to do, and then let the computer flash what it had decided to do. Then you could announce, let's say, beforehand, what the optimal value is going to be, and, if, indeed, DYGAM is doing its job, it's always going to win.

I was going to say something about the new work coming out in sequential models, but I think Bill Lucas and John Mayberry said something about that already. So, I think that's about the end of what I have to say.

Prof. Lucas: Our final panelist is John Tomlin from the Institute for Advanced Computation.

Dr. Tomlin: I am very much a novice in this area. I came here to learn, not to expound, and where I do have to expound to do so. You heard from Zack Lansdowne yesterday on the subject of implementing a DYGAM-type system on the parallel processor. I think he explained it rather well, and there are a num-

ber of questions that seemed to arouse interest so I will just make a few comments on this and perhaps try and answer any questions that might come up.

First, let me say just a few words about using large parallel processors. I think there is a sort of syndrome here with each new computational development that has technological promises. Things get vastly overblown, a few users get burned, and everyone thinks it is a bummer. This happens I think in almost any kind of computational development, but the fact is that these things, like the Illiac 4, do work now. They solve important problems that you can't solve on anything else, and for finite amounts of money. So this field (DYGAM) does appear to be a suitable one for implementation.

There are all kinds of problems where people say "Great, let's use 64 parallel processors, we can really grind this problem to dust." But this approach is just not suitable. You have got to have a data structure and a kind of program that lends itself to parallel processing and a lot of things don't. Some do, beautifully, solving differential equations, things like that where you can always do 64 things at once if you have the right grid size.

Other things don't lend themselves very well. You have 90% of your machinery sitting idle because it can't all be used at once.

Well, it seems to me that a DYGAM-type system is really rather amenable to this and we can get some kind of logarithmic improvement in problems of reasonable size, say around ten state space dimensions.

Now, as you saw from Zack's (Lansdowne) slides, large amounts of time are going into the backward phase of the program, and what is happening there is essentially just building up the data base for the forward search. So far, this is what I have concentrated on, this backward process, because all the algebra and arithmetic that is there is used over again, only in smaller quantitites in the forward search, so we can do the backward part right. We can do the forward part right if we know what we want to do.

Yesterday, a number of people — Jerry Bracken and others — raised the problem about what to do with this forward search when you only have an approximation to the payoff surfaces. One suggestion was not to do anything at all, but just to carry the strategies back with you. This may be possible if you have discrete state space points and you can sort of remember the strategies that go along with them. If you have to interpolate between them, and the surrounding points all have different strategies of where to go from there, you are in the cart unless you can do some sort of forward search. So, some sort of forward search procedure is necessary, and it's becoming fairly clear to me that this has to be refined fairly extensively before a lot of users are going to be happy with it. Just from the comments today on the panel, there have been people who have said that this sounds like a great way to sell these problems and others who have said, "well, the solutions are so unreliable, I don't give a damn." Some way is going to have to be found so that these people might give a damn.

Prof. Lucas: We have a few minutes for questions to the panel or to the speakers or general comments.

Dr. Anderson: What would be your, or your sponsor's objection, to say, building a simplified model of BALFRAM? Not an equation you can solve in closed form, but a simplified model, and then using the techniques of OPTSA or DYGAM to solve that, and then using the strategies that came out of that as inputs, as fixed strategies in BALFRAM.

Dr. Goheen: That's a good idea, and I think in light of what we learned with our experience we probably would like that idea a lot.

Dr. Anderson: There aren't objections?

Dr. Goheen: I don't think so. No. I don't think there are any.

Prof. Lucas: Jim?

Prof. Taylor: I get the feeling that I am a professional kibitzer in this business. I do have the opportunity of being institutionalized, so to speak, in a combination of both academic and DoD channels. So I would like to just make a few comments on some activities that many of you may be more familiar with than I am but that I think tie together especially nicely in this conference.

For example, Al Dobieski spoke about CATTS. Now, CATTS actually is an evolution of a Marine Corps project that the Navy Department worked on developing, and this is a system that is now called Tactical Warfare Simulation Evaluation and Analysis System (TWSEAS). The Marine Corps had two problems: one was command post exercises to get more credible assessments than just umpire judgments, and the other was to give some feedback in real time to training exercises. One of those programs is called Tactical Exercise Simulator and Evaluator (TESE). The other is called, Tactical Warfare Analysis and Evaluation System (TWAES).

The one that was the command post exercise sort of gave rise to the spinoff of CATTS. The Army looked at what the Marine Corps was doing, and it turned them on. They went out with the notion that it was nothing that money couldn't solve, and TRW did the developmental work.

Right now, I think the really exciting thing is that the Marine Corps is seriously talking about having large ground operations for training being controlled by a computerized system. In other words, the forces are deployed. The Red and Blue forces come together, and then the computer does the assessments. So, let me just cut off there.

One point of commonality is Marine Corps work. Another thing that's in the same vein is the Air Force work in instrumented combat ranges. Out at Nellis Air

Force Base there is a highly instrumented range in which high performance aircraft do their thing in a highly controlled monitored environment for the training of pilots. I saw a briefing at the May Operations Research Society Meeting, and it was really fantastic. There are big payoffs. For example, a pilot thinks he is doing evasive maneuvers. He goes behind a land mask, doesn't change his heading or course, comes out again, and he wonders why he is shot down. Well, now the pilot in postbriefing can get some feedback on what can be improved in the tactics. The same type of thing is involved, or course, in the Marine Corps landing exercise under computer control.

The Air Force is also interested in playing ground combat units in the same environment, so I think there might be some profitable interaction between the instrumented combat range people and the Army and possibly the Marine Corps.

A last element, that is similar in a sense, which was the forerunner of the Air Force work, is land combat operations. Now, many of you who are on the east coast are not as aware of the things going on on the west coast. Out there you can do things like take tanks, go out on a range, fire live rounds and have it all instrumented, and you are not going to bother anybody because there's lots of territory.

Well, out on the west coast is a facility called Hunter-Liggett Military Reservation, which is maybe 10 by 10 square miles. That is the combat laboratory for ground combat, and over a period of time, Jack Stockfisch, who is sitting over there, was one of the driving forces behind the small arms experimentation. They got some instrumentation. Then the next logical thing was to go into forces in real-time, two-sided, nonlive fire, very similar to the Air Force activities. For instrumentation you fire lasers, and you don't actually fire bullets. So, again, there's some commonality.

The Army also has another even larger facility at Fort Hood, Texas. They do the same types of things as at Hunter-Liggett, but instead of the platoon and squad level activities that are done at Hunger-Liggett, they can go out to battalion activities in a slightly less controlled environment. So, I think all of these things have some commonality between things that Al Dobieski was talking about.

As we look for the payoff for DoD in these OR techniques and computer technology, I think it may be much more likely in the area of training than in research. That is, I think commanders, operators, and participants will be put into a simulated, but realistic environment where they can be briefed after the fact and can say, "gee, I did a dumb thing, and in the real world I would have been killed." So, I think for training there's probably an even greater potential for OR. I'm only mouthing things that Ed Paxson of RAND explained to many of us at the May meeting.

If I'm not rambling too much, there is one other thing that I think should be kept in mind about Game Theory — its main function is to do what the artillery does in Army circles, to lend dignity to what otherwise would be a bloody brawl. It provides a vehicle of communication on the basic structure of the things we're talking about.

Back in 1964, NATO had a conference on the applications of game theory to military problems. The concensus among all those there was that there were none, but I don't think they're quite right. Game theory provides a very important framework for looking at these problems and for communicating the structure, the formal structure. With more interaction between the academic, think tank, and military world, I think we have to develop a clear conception of the uses for our models.

Now, the operators like to see very detailed, credible models. The researchers, to understand those detailed, credible models, want to do some parallel modeling to get simplified structures so they can communicate among the analysis community the salient points that are being incorporated in more enriched structures.

Prof. Lucas: I must say, when you referred to the artillery, that 25 years ago some people solved some duels and then said two or three smart pieces or smart bombs would be better than two hundred just thrown out not so smart. So I think that if some people had looked at that work a little more closely, they may have emphasized some of these things much earlier than they actually did. In any case, some of these ideas have been around for some time.

Dr. Huber: Well, I'd suggest that even though many of these devices may be conceived for training they are actually quite good research tools. I have been thinking, after the problems that were discussed yesterday, that software makes all the difference to, say, the question of how to employ weapons. I think that many of these devices have a great potential in really asking the question, "how should you really employ your weapons?" "What kind of doctrines should you adopt?"

I remember that Wilbur Payne, in the 1974 conference, on land war modeling, which took place in Germany, gave a paper in which he exploited just this idea. He said, "well, we do a lot, we invest a lot of money into hardware developments, and we develop new weapons, and we only find out how to use them after we get them to the troops and start developing doctrines for things that exist. Why don't we do this a priori?" It probably would prevent us from going all three ways in weapons development, and we would

even have the possibility of training people long before the weapons are in inventory so that you could actually greatly abbreviate this turnover phase or the changeover phase from one weapon system to another.

Mr. Louer: The last point Dr. Grotte made reminded me that when this conference first started one of the first things that I said was that in the Army, we are applying theater models to our force-planning problems, and one of the difficulties in doing this is in measuring the effectiveness of alternatives. We don't have a good solution to this. The example (Slide D), which was shown of one side aiming for a two-day war, and the other side aiming for a seven-day war, says this strange thing that the Blue won because he was in better shape at the end of seven days. Now, I can give you some real-life examples to that problem without getting into any classified information.

The Army has looked at some alternative forces against the Red threat in Europe and they played the game out for 180 days. Starting the gaming with one force configuration, they find that it loses badly. It is swept off the continent even before 180 days expire.

Another force, then, is given some additional help, and at the end of 180 days it is in great shape. This force is sitting with very little area lost, and the Red force is depleted, but Blue is in good shape. However, if you look carefully, and the unfortunate thing is that the decision is made that favors this seccond force, but if you go back and look carefully at the history of the battle, you'll find that back around D+10 or D+15, Blue force has maybe only 10 or 20% of his tanks surviving at that time. His personnel are down 30 to 40% in both games. Yet, at the end of the war, in the second case, Blue is in good shape. But the thing that is important that we have failed to look at the fact that if you really put yourself in that D+15 situation you don't have many tanks, your personnel are low, and the Red force is continuing to build up. You would have never continued the war.

So, what I am getting at is that we need some better way of measuring the effectiveness of forces, and this is a good example of the case that if you only look at the end result (after 180 days), that is not the correct thing to do. That is where I say we need help. We need to work on that problem.

Dr. Huber: I would just like to ask one question. Would you think, even though you may not be very sure of what to do in terms of a decision in those two cases, that the very fact that you have the information about both force alternatives was to your advantage?

Mr. Louer: Yes, that's right. What I am getting at is the information that was being developed in the model somehow failed to get distilled into the decision. That's where the problem is.

Dr. Huber: Okay, Fine. So, this does not necessarily detract from the utility of having some kind of decision modeled?

Mr. Louer: Communication is the problem in presenting results of the battle. That's . . .

Dr. Huber: I am personally not too uneasy that we are not able to come up with the one overall measure of effectiveness. Because, I think, if we could, we would just kill so much information that the decision maker needs to consider.

Mr. Spaulding: I don't think the decision maker wants one damn measure. He wants to know what's going to happen. He wants something that is small enough in information content to be comprehensible. But he wants something on which he can exercise his intellect and his judgment and his knowledge of features in our model so that he can say, "well, the model gives me this, but I know there are other things." And he can exercise his judgment on this comprehensible set of information. When you try to tell him what the optimal answer is, you are preempting his perrogative, and so you have to be very careful, I think, about trying to provide the decision maker with optimal solutions because they're optimal in this fictional game world but they're not optimal in the real world.

Mr. Louer: They may be optimal in the real world, too, but you still better be damn careful. He doesn't like to . . .

Mr. Spaulding: He doesn't like to be preempted. He likes to exercise his judgment.

Prof. Lucas: A friend of mine working at Mathematica told me they did a problem of redistricting for the state legislature in New Jersey. Mathematica apparently came in and gave them an optimal solution; just one solution, just one optimal. But the legislative folks said, "well, maybe we would like to make a choice, we would like to see some of the near optimal solutions. So the next meeting they have a great big pile of computer printout and they looked up and down, and up and down, and up and down, picked the second page, and said "we like this one."

Now Martin Shubik made some references to work on duels. I think the person I would go to who may be the most current on literature on duels, at least unclassified literature, is George Kimeldorf — I am not sure if I spelled this right — at the University of Texas, at Dallas. I think he is still active in this area, and probably fairly current on the literature, at least unclassified.

I would like to thank the speakers of this session and the panel members and we are almost on time, and we should go onto the next session.

Session V: Data Base Requirements for Theater-Level Models and Data Related Data Input Generation Problems

Opening Remarks

ROBERT SCHNEIDER
OSD, PA&E

Mr. Schneider: Hopefully, we'll get to the rather controversial part of the conference. We have been leading up to it now for the last two or three days, and every once in a while we hear the problem of data bases and data — we don't have data. I am hopeful that we will be able to generate quite some discussion about this during our remaining time.

Incidentally, I am Bob Schneider, from the Assistant Secretary of Defense, Program Analysis and Evaluation. I say that slowly because PA&E gets called many things, but its real name is Program Analysis and Evaluation. I was rather disappointed by the fact that as we were building up a thoughful and healthy discussion here yesterday, we ended up getting cut off because of time. So, I am hoping that we won't have to say too many things; say, we spend an hour or so talking about the problem, and then spend more time in the discussion later on.

Perhaps a large part of the problem that we have with data bases is indicated just by the fact that I am the chairman of the data base session because in many ways, I am an outsider looking in. First, my major in college was mathematics, not operations research or systems analysis or anything like that.

My graduate work was still mathematics, getting a smattering of OR techniques, and so my outlook on life typically is one of a skeptic when I look at problems unless I can see very clearly the logic that is involved and that all of the steps and operations are built on some sort of logic. I am also formally from the Army, and while I was on the Army staff working with operations research and with large problems, again I was kind of skeptical because it was my job to make sure that whatever the contractor came out with, while it may have been theoretically very nice, was practical, something we could use. I think the third reason I end up being a skeptic is because I didn't start out in theater combat models but in force structure models at the theater level. And in developing a force structure for a theater, you end up dealing with a whole bunch of people who are making many, many assumptions, and are very quick to pick out numbers not based on some sort of logic, but because that's how a particular proponent thinks he can get what he wants.

The engineers, why do you have as many engineers in the Army? Well, I claim it's because that's how many engineers we can get. Artillery is the same way. How much artillery does a commander want? All he can get. That's the typical answer, so when it moved over into the combat area, I began asking the same questions. What is all of this built on? Where are all of these numbers that we're getting? When we make the assumptions, what is the logic behind it or what is the empirical data and that sort of thing. I was very much surprised to find out that we have some very large assumptions that we just kind of blandly go along with to get into what I see as the shiny toy. So, I see two big problems with the data, and with modeling, as it's associated with the data.

The first problem is that the model development is decades ahead of data development. So we deal so often with the typical lamp-post problem. I think most of your are familiar with that, where the guy is looking for his keys underneath the lamp post when in fact his car is down the street some ways, and he has never looked down that street because it's dark down there. So what we do is we continue to look under that lamp, even though the basic problem, and the likely solution, is down there where it's dark, because it's easy to do and it's something that we can continue to work on.

I think that is a very real problem in the world today, and I don't blame the model builders because the model builders, as are most of us, are motivated by making the buck. And, as long as we can get money, as long as people will pay money to do this, then great, that's what we'll end up doing, even though we may recognize that's not where the problem is. So it's not really the model developers, I don't think, that are as much to blame for this as it is those of us who are in DoD, who are spending the money. So, I don't put the

blame on many of you that are here, as much as on myself and others of us.

We have this very real problem in data development because of our system of competition. I alluded to this on the very first day. We have a bunch of proponents who are in essence selling a product, whether it's the Army selling a product versus the Air Force, or whatever, to gain money within the OSD or DoD Arena. Or take us, for example, we turn our coats halfway through the year; during the first half of the fiscal year, we're trying to figure out between the various services which should get the money and then we end up the later half of the year turning right around and saying, "now, let's use some of those arguments because we want to get money from Congress." So we end up in many ways, in a sense, being two-faced; we're forced into that position. But, because we've got this system of proponency, the tendency then, or at least the temptation, is to say "I'm not really so worried about all the stuff that's in that model. As long as the answer comes out right, I'll take it." When we think in terms of models developing requirements, it's my contention that as so many of you have brought out, there are too many uncertainties in the development of models to be able to use them in a pure requirements mode. I think it's been said many times that these models are used to gain insights. You can look at things within the models that you feel fairly comfortable with, and you can compare different weapon systems, for example, in some of the models where you understand everything that's going on — that sort of thing. But when you get into the requirements mode, you have a serious problem. Yet, within our system of competition, that is exactly what we use models for very often. What's the estimate that you're going to come out with, what's the requirement that you are going to come out with? So, I see difficulty in the identification of where the real problem is and that is in data development because it may be considered counter-productive in this environment I've been describing. It's kind of nice sometimes to hide things.

Okay. Another problem is that of data-gathering. We are not set up to get the kind of data that models need. I have been working with DIA for quite some time in data gathering, and I think that their problem is typical of any intelligence agency. Anybody that is trying to gather data about somebody that isn't going to tell him what the data are has a problem. We are so technically oriented, and, as a matter of fact, technically capable of gathering data from above that what we end up getting, for the most part, is total numbers of things in a certain area. It's what we can see from above that we feel fairly comfortable about. So if you ask somebody, an intelligence analyst, tell me how many tanks there are in a particular area, or tell me even what kinds of weapons there are in particular area of the globe, he feels fairly comfortable with that. As a matter of fact, he even feels kind of comfortable with saying, "well, that's kind of a battalion-sized unit, or that's a division-sized unit," or whatever, and in a general area. But if you ask him, "Now, I want you to tell me how is that tank battalion or that motorized rifle regiment organized, how is it put together, what are the tactics going to be, how is it going to fight, how is it going to use those weapons?" They just go ape because they haven't the slightest idea.

I think in model development we're going in just the opposite direction that data development is going in the sense that where more data are developed, we see a problem of demanding from the analyst something he can't do. That is why I was heartened by the discussion here of people saying, "look for better or for worse we really do need simpler models, not more complex." Seth Bonder was saying that the community has become schizophrenic, and I say, well, that's no justification for him becoming paranoid. The thing is, it's true, we are schizophrenic, in that we're going in these different directions, but that doesn't mean build a bigger and more complex model that you can't get data for.

Now, I know people are saying, "yeah, you can get the data," and, hopefully, we'll discuss that a little bit. I would hope that the various panel members would be able to give us some insights as to how to solve this problem, because so far all I am telling you is that there is a problem. I have some ideas, and I'll save them for later on. But I really believe that no matter how much we discuss the niceties of the theory of gaming and all the things that we can do in gaming it is still way too far ahead of the very basic need for gathering data that can be used, and that we have got to spend more time on that problem.

There is a key question here. Most of the people I see here are professors and doctors and so forth, and then you come down to a Mr. Schneider. The key question is how do we get the academic community and the people who spend the money for developing models turned around to finding ways of making the data applicable to all of this?

In this same vein, another thing I might say that I am heartened by is that Larry Low decided we should have a session on data bases, so, with that, I would like to ask Ray Bednarsky to give us a presentation on what is being done now within the Department of Defense to get a handle on this problem, by no means the complete solution, but an attempt to solve it.

Ray Bednarsky is from Command and Control Technical Center, and is going to talk to us about the Department of Defense Force Planning Data Base, which I have been involved with, and also, with a data base that I think you are much more interested in that is being developed for war games for SAGA.

28 — *Overview of Data Base Problems*

MAJ. RAYMOND BEDNARSKY, USMC
Command and Control Technical Center

Surprising and significant advances in the development and management of data bases for theater models

Maj. Bednarsky: I think we have come to the point we have all been waiting for in the program. I represent the last speaker. If you will give me just a movement to get set up, we'll discuss data base overview.

Perhaps, I can do something that Bob (Schneider) alluded to. If I'm successful here, maybe I can move that lamp post a little bit further toward the dark area, at least I will try.

Thus far in the conference we seemed to have talked a great deal about conceptual designs of war games, that is the processes that are represented in the games, the way in which they are represented and their level of detail. These, of course, determine the input data requirements for the games, and underlie most of the monumental problems associated with the preparation of the input data. Therefore, I would like to present a brief overview of the data base problems encountered by theater-level modelers as they attempt to employ the state of the art in models.

What I would like to do is to restrict the scope of the overview to the data base requirements and input generation problems. I would like to present to you a statement of the problem (Slide 28-1). I would like to talk a little about the current data-gathering process. Then, I'd like to discuss the high and low level data needs for theater-level models, and then propose to you a data base method, so I come with some sort of a solution in hand.

Slide 28-1 OUTLINE OF TOPICS

1. Statement of Problem
2. Current Data Gathering Process
3. High/Low Level Data Needs
4. Data Base Method
5. DoD Force Planning Data Base
6. OJCS Weapon Characteristics & Performance Data Base
7. Technical Problems
8. Management Problems
9. Recommendations
10. Conclusions

I think it's appropriate that we look at the two major data base endeavors in the Department of Defense: the DoD Force Planning Data Base, and what we call the OJCS Weapon Characteristics and Performance Data Base. We should discuss a few of the technical problems involved, which are monumental, and some of the management problems; I then come with complete recommendations and conclusions.

I'll put on this statement of the problem (Slide 28-2). I should tell you at the outset, that my perception of the problem is somewhat shared by my responsibility at the Command and Control Technical Center to convert, document, modify, maintain, and operate 22 theater-level models for the Studies, Analyses and Gaming Agency of the Joint Chiefs of Staff. Now, I believe that a great amount of attention has been focused on the research and development of conventional nuclear and chemical theater-level models and rightly so, because that's where the big bucks are. However, very little attention has been given to how the aggregated,

often ill-defined model input parameters relate to the available intelligence data. I often find that little documentation exists describing the transformation from low to high level model input data, and, in practice, a wide variety of ad hoc aggregation techniques were used to generate the data.

SLIDE 28-2 STATEMENT OF PROBLEM

- CURRENT THEATER-LEVEL MODELS REQUIRE 6 TO 9 MONTHS FOR INITIAL INPUT PREPARATION
- MODELS BEING ACQUIRED OR PLANNED FOR FUTURE USE MAY REQUIRE TWICE AS MUCH INPUT DATA
- SERIOUS QUESTIONS HAVE BEEN RAISED ABOUT THE VALIDITY OF DATA GATHERED USING THE CURRENT PROCESS, BOTH AS TO USE OF BEST SOURCES AND AS TO MEETING THE ASSUMPTIONS IMPLICIT IN THE MODEL BEING USED

If the employment of high resolution models requires the generation and preparation of huge quantities of complex input data, concerned with weapons and other system characteristics, and performance capabilities, such as environment you know, weather, terrain, logistics, order of battle and operational doctrine and tactics, then we must recognize the need for a centralized, comprehensive data base in the Department of Defense under the management of an able data base administrator, and, finally, the problem is one of data effort redundancy.

Many of us in DoD are expending a large number of dollars to produce the same number. I have here an example, very appropriately so. In this slide (Slide 28-3), I would like to compare VECTOR 2 with other models. These are the state of the art models, gentlemen; we have here represented the OSD model, the Army's main model, and the Joint Chiefs of Staff's main model for at least the next few years.

If you notice, VECTOR 2 has what we tried to achieve for a long time; a renovation in the state of the art in that we have very exquisite attributes. These cost us something in comparison with other models, and, naturally, the cost is that of increased amounts of data input.

If we look at these attributes, which take up over two slides, we find that, in many cases, the VECTOR 2 model that has been developed along with VECTOR 1 really presents to us a nightmare from the input standpoint. Before we can crank either of these models, we need to really get busy and accumulate possible scenario data.

I thought it was very appropriate when the gentleman asked Seth Bonder, where do you get the data for that model? We have been asking him that since about 1972. Well, we think that working together would be all right, working together with other agencies in the Department of Defense. We think we have a solution to this problem.

I would like to show you, if I may, the current method of using a model (Slide 28-4). We simply take the data, input it into the model, run the model, and get the answer. This is what the contractor would have us believe. However, that is not the case. It is rather a simplistic view.

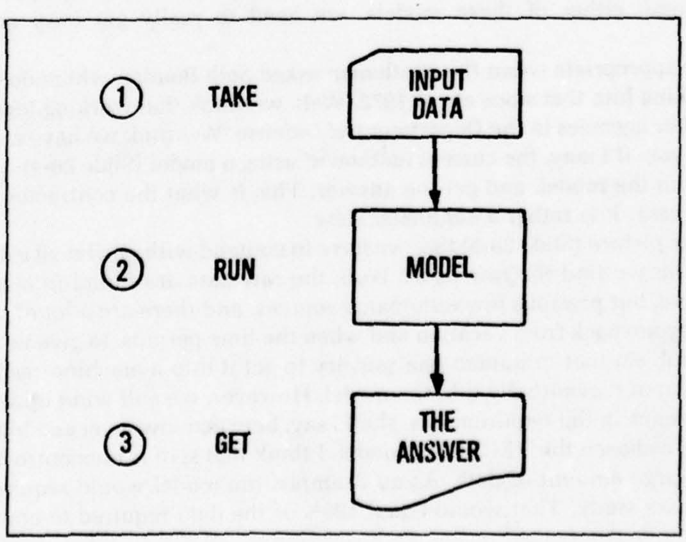
This is a more realistic picture (Slide 28-5) that we have to contend with on almost a daily basis in the Department of Defense. Where do we find the raw data? Well, the raw data are found in classified and unclassified documents. There are some, but precious few automated sources, and there are a lot of very subjective analysts who are willing, when they are back from vacation and when the time permits, to give us their guesstimate of the data. We then take that, subject that to human analysis, try to get it into a machine-readable format within the time permitted, and then input it eventually into the model. However, we still wind up with only low level data.

Let's look at the difference in the requirements, shall I say, between low level and high level data. As a vehicle to discuss this point, I've chosen the VECTOR 1 model. I think that spot is noncontroversial so far as we have gotten but yet requires a large amount of data. As an example, the model would require approximately 20,000 values per scenario on a case study. That would equal 100% of the data required to crank the model.

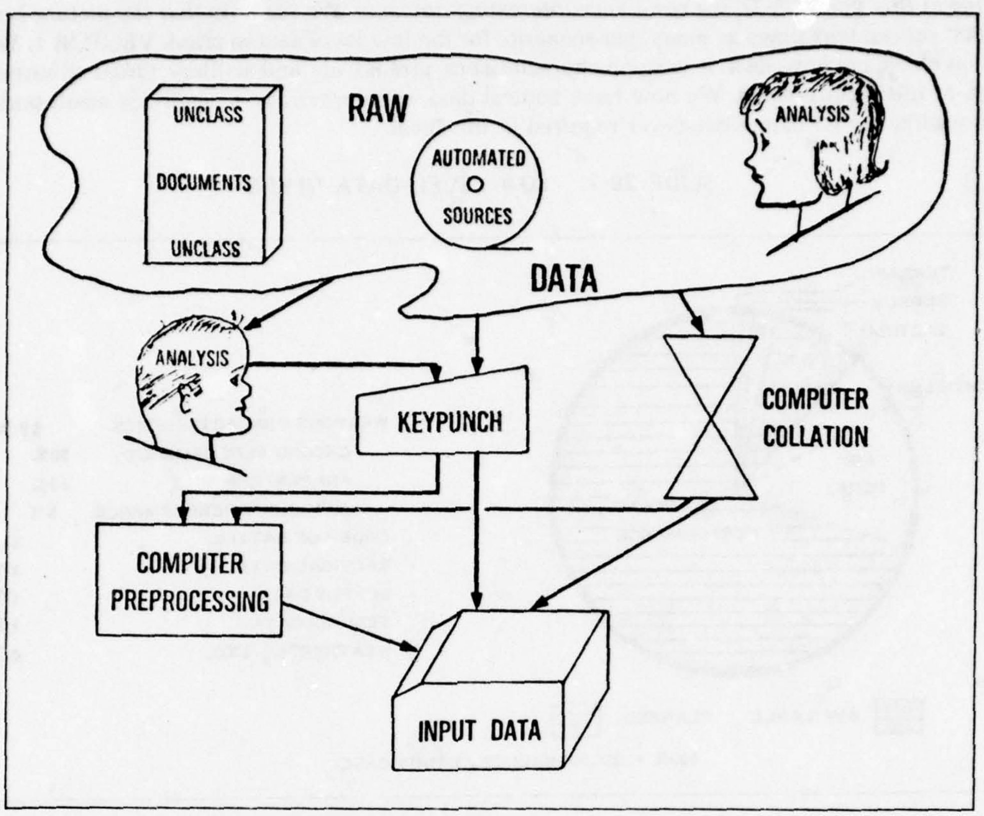
Slide 28-3 — COMPARISON OF VECTOR-2 WITH OTHER MODELS

Attribute	VECTOR-2	VECTOR-1	IDAGAM I	TACWAR	LULEJIAN V	CEM IV
Types of weapons	28	21	22	22	37	50
Types of supply	34	27	1	4	2	5
Dimension-changer	Yes	No	No	No	No	No
Non-zero FEBA width	Yes	No	No	No	No	No
X/Y location of ground forces	Yes	No	No	No	No	Yes
X/Y location of air forces	Yes	No	No	No	No	No
Multiple terrain across sector	Yes	No	No	No	Yes	Yes
Weather	Yes	No	No	Yes	No	No
FEBA movement on objectives	Yes	No	No	No	No	No
Movement across combat arenas	Yes	No	No	No	No	No
Time (number of clocks)	6	2	1	4	2	1
Multiple activities per day per unit	Yes	No	No	Yes	No	No
Command and control	Yes	Yes	No	No	No	Yes
Modular user-specified tactics	Yes	Yes	No	No	No	No
Communications	Yes	No	No	No	No	No
Command heirarchy	Yes	No	No	No	No	Yes
Intelligence	Yes	No	No	No	Yes	No
Target acquisition	Yes	Yes	No	Yes	Yes	No
Construction	Yes	No	No	No	No	No
Maintenance	Yes	Yes	Yes	Yes	No	Yes
Post-processor	Yes	No	Yes	No	No	Yes

SLIDE 28-4 CURRENT METHOD OF USING A MODEL

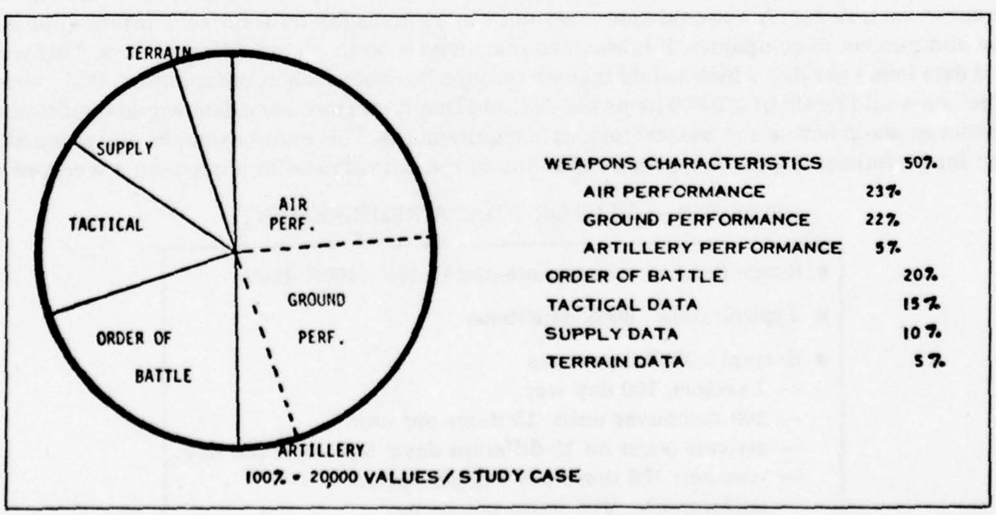


SLIDE 28-5 DATA GATHERING PROCESS



As you can see by the pie chart (Slide 28-6), the weapons characteristics portion consisting of air, ground, and artillery, make up about 50% of the model requirements, order of battle, approximately, 20%, and, then, of course, we have our tactical, our supply, and our terrain. That doesn't seem to be that great of a problem yet. Only half being the weapon characteristics portion, and, of course, we do have a method of dealing fairly well, thanks to Bob and his crew, with the order of battle.

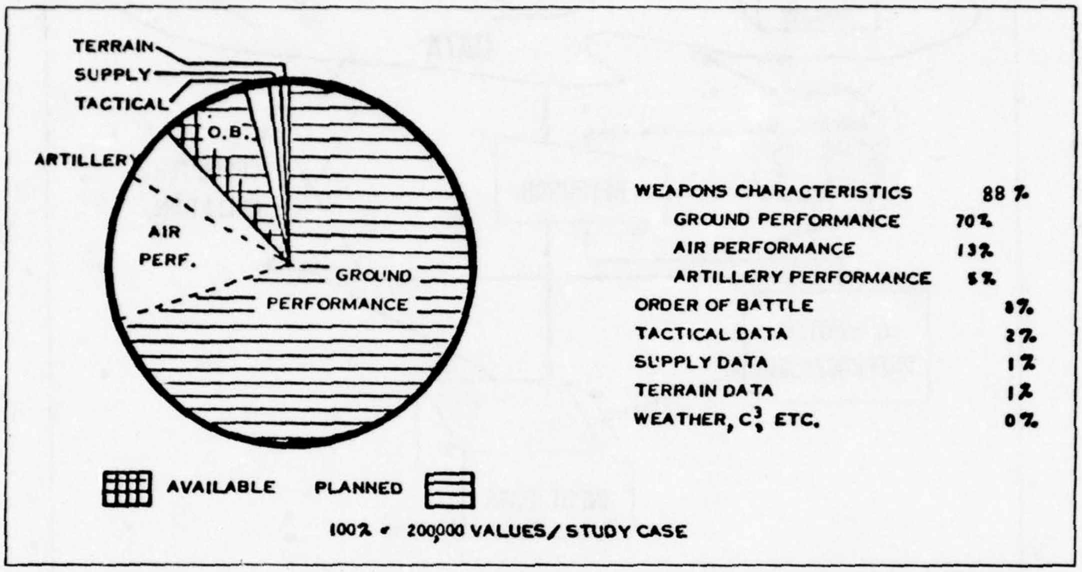
SLIDE 28-6 HIGH LEVEL DATA NEEDS



Well, that isn't the whole game because I define high level data as that required as the input to a model. Before we can get this aggregated or high level data, we must have the raw low level data.

Looking at this (Slide 28-7), we see a very interesting outcome. We see now that the picture has changed. I need 200,000 values, tens times as many, per scenario, for the low level data to crank VECTOR 1. We see that a great portion of the pie now, 88%, is weapon characteristics, ground, air, and artillery. Order of battle happens to be only 8% of the requirements. We now have tactical data, supply, terrain, a relatively small portion, and, of course, no weather or C^3 data whatsoever required at this level.

SLIDE 28-7 LOW LEVEL DATA NEEDS



We have these data available, for the most part. This is the plan, all of this. If both of these things come to fruition, we'll then have data required to run the model. Now, can you imagine getting out 200,000 pieces of machine-readable data and then subjecting it to aggregation by hand? The problems are monumental.

Well, that's VECTOR 1. Let's look at the next evolution in the VECTOR series, which provides a good vehicle, VECTOR 2. Emphasizing the importance of day-to-day development for VECTOR 2, Slide 28-8 shows VECTOR 2 input data requirements in terms of items in numeric value. Data requirements for a run are likely to range from 14,000 for a small test case to a maximum of 400,000 items. Now, a typical operational run will likely require 200,000 to 300,000 items.

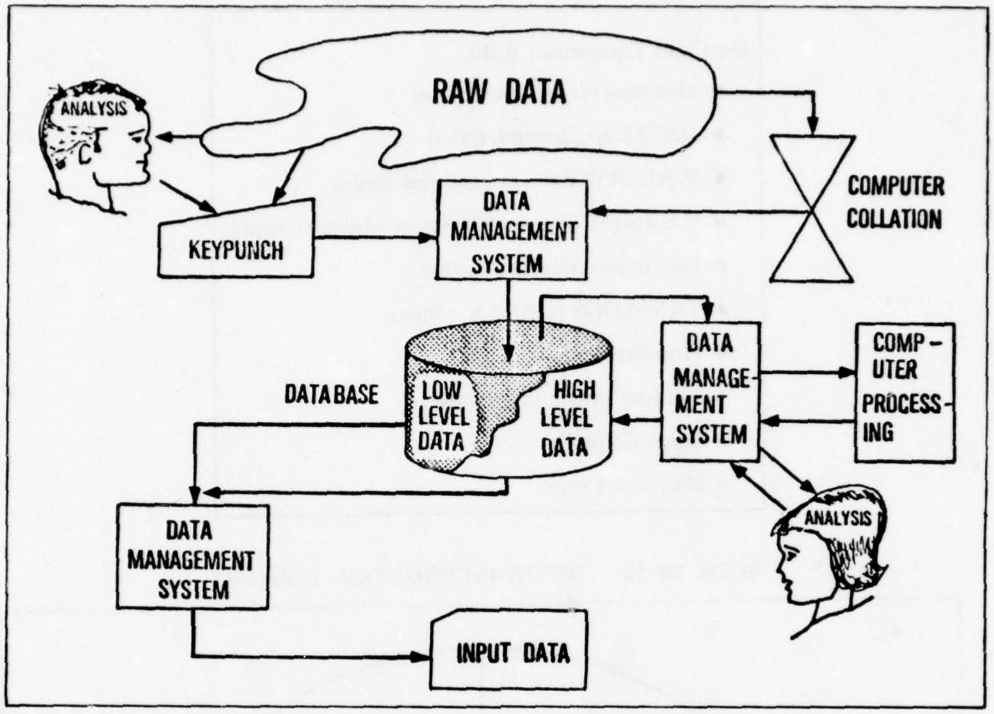
As an example of a very large requirement of 380,000 items, 7 sectors, and 100-day war is shown, with everything different in each sector. A hundred maneuver units at 15 items per unit, for such information as location, zone, nation and number of companies. It is assumed that arrivals occur on ten different days. This would result in about 500 data items per day which would include resource arrivals such as ammunition, POL, weapons, and sensors. Weather would be about 170,000 items per day, and finally, performance data would require about 31,000 items per sector or about 60% of the 380,000 total data requirements. This entire example, and especially this last requirement for performance data, is a strong indication of the critical need to implement a weapons data base.

Slide 28-8 — VECTOR-2 DATA REQUIREMENTS

- Range (total number of data items): 14K - 400K items
- Typical: 200K - 300K data items
- Example: 380K data items
  - — 7 sectors, 100 day war
  - — 100 maneuver units: 15 items per unit
  - — arrivals occur on 10 different days: 500 items per day
  - — weather: 170 items per day per sector
  - — performance: 31K items per sector

In the data base method that I propose for handling this problem (Slide 28-9), we would again be left with the problem of accumulating, finding the sources of raw data first, and assuming, if you will, that we could find these sources of data, we would somehow subject them to analysis and produce a machine-readable piece of data. Then through a data management system, we would input into a data base, which would contain at a minimum, the low level data as well as the high level data.

SLIDE 28-9 DATABASE METHOD FOR GATHERING DATA



Now, low level data would perhaps be aggregated by means of models and preprocessors, through a data management system and through some computer processing, and, naturally, human analytical intervention, and would then be stored as high level data within this data base.

We'll get more into detail in the conceptual process as we go along.

Then, theoretically, we would be able to select, either low level, high level data, or a combination thereof, through the data management system, and input the data into the model. Again this is a theoretical solution.

For a moment, let's digress, and look at part of what we have to get in the way of data — this is the Department of Defense force planning data base on NATO and WARSAW pact forces. This is what Bob and Mark Liddy and several others here worked so hard for, and we now find such a convenience in the area of modeling.

I am sure many of you recall well that this orginated with the formal title of NTFAM 3, which was the NATO Task Force Action Memorandum of 16 August of 1973 (Slide 29-10), in which DP&E was tasked to develop a data base of NATO and WARSAW Pact conventional forces for economic information. Then the worst possible thing to do was to update it on a continuing basis. They are still engaged in that.

Slide 28-10 — NATO TASK FORCE ACTION MEMORANDUM NO. 3 - 16 AUGUST 1973

- Task: Data base of NATO and WP conventional forces and economic information
  Support DoD, NTF, MBFR and NATO fact book
- Responsibility: DP&E
- Coordination with: ASD(I), JCS, DIA, CIA, ISA, ASD(C) and DDR&E
- Timing: Updated on a continuing basis

In the DoD Force Planning data base, we now have these data parts (Slide 28-11). You notice that primarily it's an order of battle, featuring unit and equipment. We have some transportation logistics, routes and nodes. We have some ammunition, equipment standardization, scenario plots, and a world outline. There's a lot of work that is being done on this.

Mr. Anderson: This is for current, not 1989 or something, right?

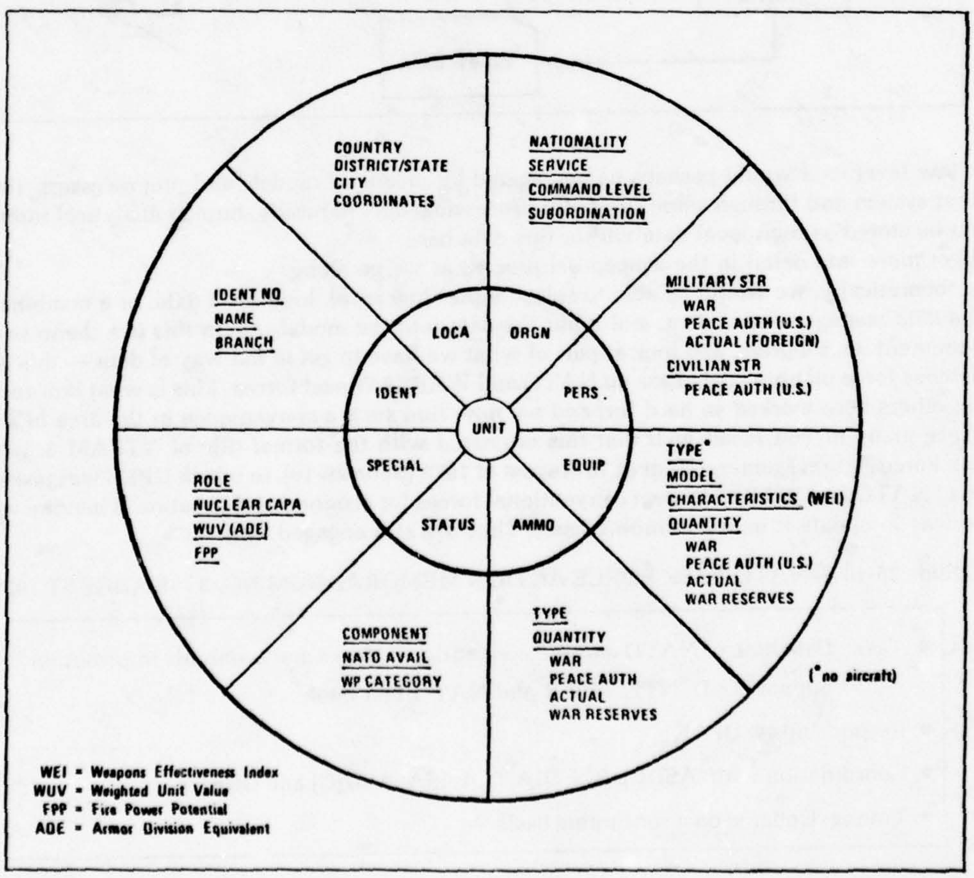
Maj. Bednarsky: This is for current.

Slide 28-11 — DATA FILES

Unit and Equipment (OB)

- Mid-east air, ground, naval
- NATO, air, ground, naval
- WARSAW Pact, air, ground, naval
- U.S. forces (Army, Air Force, Marine Corps)
- Equipment characteristics
- TRNSLOGS (routes & nodes)
- Ammunition
- Equipment standardization
- Scenario plots
- World outline

SLIDE 28-12 DFPDB INFORMATION CONTENT



This slide (Slide 28-12) is background for anyone who has not used this data base material. It's very well thought out. It's all related and keyed by unit. The information, again, features the unit as a central segment and then from that, associated equipment, ammunition, personnel, origin, location, identifications, special characteristics and status are determined. For instance, associated with certain units are their equipment by type, model, characteristics, quantities, etc. We can also look at the ammunition portion of it, and we can see that we have the type and the quantity — for war, peacetime authorized, actual war reserves. Looking at the status, then, of the unit, we can see the component, the NATO availability and its WARSAW Pact category. Special characteristics would be noted such as its role, whether it has a nuclear capability, whether it has an armor division equivalent, its WUV, and its fire power potential. Again, we have its personnel where we feature the military strength — for war, peacetime, actual civilian strength associated with that unit. In addition, we have a unique identifications, countries, nationality, command-level subordination. It's a fairly well-structured, well-thought out piece of work, and of immeasureable aid to modelers, and force planners.

Let's look at some of evolutionary process that has taken place to get where we are today in the DoD Force Planning Data Base (Slide 28-13). Here's where the blood, sweat and tears came in. The data started to be inputted in the data base in 1974, and you can see that we've done a lot of work through 1977, and have some work to do in 1978, and, of course, as you well know, the data base, will never be complete. There will always be enhancements to it.

Slide 28-13 — IMPLEMENTATION YEAR FOR CURRENT DATA

	Unit Info		Unit Manpower			Unit Equipment			Non-Unit	
	Act.	Res.	War	Peace Auth	Peace Actual	War	Peach Auth	O/H	Equip	Pers
Forces U.S.										
Army	74	74	74	74	78	74	76	74	77	77
Marine Corps	74	74	77	74	78	77	74	77	77	77
Air Force	74	76	77	74	78	77	77	77	77	77
Navy	77	77	77	77	78	77	77	77	78	77
NATO/WARSAW Pact										
Ground	75		75		75	75		75	77	77
Air	75		77		77	77		75	77	77
Naval	77		77		77	77		77	77	77
Mideast	77		77		77	77		77	77	77

Let's look briefly for a moment at an example of the type of information that we can draw from the DoD Force Planning Data Base, as modelers. Here (Slide 28-14) we have an unclassified, hypothetical example of the

28-14 WAR ORDER OF BATTLE (HYPOTHETICAL) — U.S. ARMY

COMMAND LEVEL	UNIT NAME	UNIT TYPE	UNIT CODE	UNIT ID	PEACETIME LOCATION	MIL PERS WAR	MIL PERS PCE AUTH	CIV PER	AVAIL	PARA MIL	NUC CAPA
4	1 CORPS			00008					AI		
2	1 ARM CAV RGT	CAV		18011					AI		
0	1 TRP AIR CAV	ARM	17058H	G2BAA	GE FULDA	200	190		AI		
1	1/1 SQ CAVALRY	ARM	17055H	G2CCA	GE FULDA	900	800		AI		N
1	2/1 SQ CAVALRY	ARM	17055H	G2DAA	GE B KIS	900	800		AI		N
3	1 ARM DIV	ARM		02003	GE				AI		
0	1 HHC DIV	ARM	17004H	AEKAA	GE FRANK	200	180	25	AI		
0	1 AVG HHC DIV	ARM		AEK99	GE FRANK	50	40		AI		
0	1 CO AVIATION	ARM	17087H	AFDAA	GE FRANK	100	80		AI		
0	10 CO DIV	INT	30017H	BV9AA	GE FRANK	80	60		AI		
1	1 BN	ENG	05145H	AFUAA	GE HANAU	900	900		AI		N
1	12 BN	SIG	11035H	AFBAA	GE FRANK	600	500		AI		
1	1/23 ADA BN	SAM	44325H	AXAAA	US BUDIN	600	600		AI		N
2	1 BDE 1 ARM			AELFF	GE				AI		
0	1/1 HHC BDE	ARM	17042H	AELAA	GE KIRCH	100	95	10	AI		
1	1/60 BN TANK	ARM	17035H	AEOAA	GE KIRCH	600	500		AI		

SAMPLE DATA

U.S. Army war order of battle. We show reporting echelons of command, command hierarchy. We've described the unit's name, the type of unit, we have a unique unit code and identification, and, its peacetime location. It's military personnel are described in terms of war authorized, peacetime authorized, associated civilian personnel, its availability, required delivery date, and its paramilitary or its nuclear capability if it has one.

On the equipment side (Slide 28-15), we can draw selected unit information such as personnel, major equipment, and weapon scores, and we can describe personnel and, basically, we'll look at the equipment. That happens to be very valuable to modelers, and we see that associated with the equipment is the WUV, the WEI, the fire power potential, and the armor division equivalents.

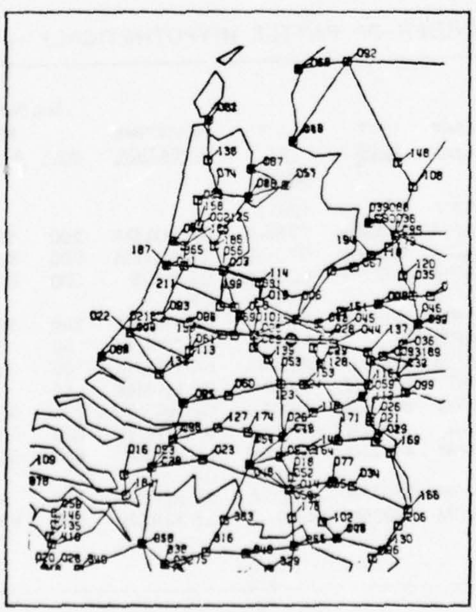
28-15 U.S. ARMY LEVEL 1 – SELECTED UNIT INFORMATION — PERSONNEL, MAJOR EQUIPMENT AND WEAPONS SCORES (HYPOTHETICAL)

UNIT CODE - 17035H UNIT NAME - 1/1 BN TANK SAMPLE DATA
PERSONNEL - WAR 500 PEACE 400 CIVILIAN 0

EQUIPMENT -	EQUIPMENT CODE	WEI CATEGORY	EQUIPMENT NAME	WAR REQ STRENGTH	WAR EFF STRENGTH	PEACE AUTH STRENGTH	ACTIVE STRENGTH
	DRB	1	5.56 M16 RIFLE	300		250	250
	EPO	2	M113A1 50 CAL	8	8	8	8
	GMH	3	M60A1 105 GUN	54	54	54	55
	HHG	5	TOW-M113	4	4	4	4
	MSE	7	107 MTR M106	4	4	4	4
TOTAL WEAPONS SCORES BY WEI CATEGORY							
		CAT 1		460.19	103.97	460.19	487.82
		CAT 2		14.30	14.30	14.30	24.56
		CAT 3		54.00	54.00	54.00	56.00
		CAT 5		4.30	4.30		
		CAT 7		3.96	3.96	3.96	3.96
WEIGHTED UNIT VALUE (WUV)				4178.329	3786.487	4021.379	4258.342
FIREPOWER POTENTIAL (FPP)				1041319	1038063	1012947	1060389
ARMORED DIVISION EQUIVALENTS USING WUV				0.06912	0.08584	0.06652	0.07044
ARMORED DIVISION EQUIVALENTS USING FPP				0.11918	0.12636	0.11593	0.12136

One last example of what is contained in the files of the DoD Force Planning Data Base happens to be this hypothetical, again, example (Slide 28-16), of the rail network in the Netherlands. Here you see the railway depicted in terms of lines connected by nodes and we have here a code, a nodel code that's associated with it. These, I grant you, would be used more by force planners than it would be by modelers, but nonetheless it exists and it happens to contain some very valuable information.

SLIDE 28-16 RAIL NETWORK IN THE NETHERLANDS



Well, that brings up the overall use of the DoD Force Planning Data Base (Slide 28-17). What is it used for? Primarily, it is used by force planners and war gamers for studies and analyses, and it also serves as a U.S. input, as you know, for the NATO Force Planning Data Base.

Slide 28-17 — USE OF DATA BASE

Users	Uses
Secretary of Defense	MBFR
	Posture Statement
	Congressional Inquiries
NATO	NATO Force Planning Data Base
Joint Chiefs of Staff	
Services	Studies
Defense Agencies	War Games
Defense Contractors	

This is a brief overview of the DoD Force Planning Data Base. Again, a monumental, very expensive effort has been undertaken by some rather enlightened people, and they also have on the negative side problems that they're dealing with even at this time (Slide 28-18). You can imagine the huge quantity of data that is involved, and the problems that it might create. Quality control happens to be one of the major problems — just getting quality data from the services, or shall I say the sources, which includes the intelligence community, of course. Then trying to plug the holes in the dyke of missing information is monumental.

Slide 28-18 — PROBLEMS

- Huge quantity of data
- Little quality control of inputs
- Missing information
- No common counting or estimating methodology
- Data not updated simultaneously
- No common equipment description methodology
- Some data sources not adequately used
- DIOBS not flexible or efficient
- Sources conflict — institutional bias
- Multiple codes
- Multiple formats

We also have a problem in that there is no common counting or estimating methodology, and there is very little common equipment description methodology in the intelligence community. You can imagine the types of guesstimates we get in that area.

And, then there is another problem — not having a standardized reporting procedure and reporting dates, the data cannot be updated simultaneously. Some of it's on a catch-as-catch-can basis, and, as a result, after we have updated certain data and get future submissions, it's an iterative process.

Well, I'd like to comment on the item "sources conflict" (Slide 28-18). The sources do conflict in many areas. In many cases it is by design and because of institutional bias. In some cases, the sources just conflict or are withheld, and you probably know more reasons why than I do.

We still have the problem of multiple codes being used by different source agencies. We'd like to get a standardization. We've come a long way on that, and I think that problem will be solved in the near future.

Reporting in multiple formats is another problem that we're working very closely on with the sources and hope that that will be solved in the near future.

Well, with that, I'd like to leave it right here, and after lunch we'll come back and we'll take a look at what I call the solution for the generalized data base problem and I'll propose that to you and we can discuss that. (Lunch break)

Maj. Bednarsky: Okay. I think I'll go ahead and get started then, as long as we have a quorum.

There is one question that I would like to answer that was asked, and I interpreted the question wrong. Bruce

(Anderson) asked if we have future data in the DoD Force Planning Data Base. We do have some future data. We have Army future data and Air Force future data that is being expanded on, so there is future data within the DoD Force Planning Data Base.

Now, to continue with what I said before lunch, and lunch by the way was very interesting because I found that many people have the same problem no matter what your level of data base generation. As I promised before lunch, I will propose to you a solution to the theater-level model data base problem. The solution I think is going to lie in what I refer to in the pie slide in which I used VECTOR as the vehicle (Slide 28-7), and I said that the problem is a lack of low-level weapons data. We saw that about 88% of the input requirements to VECTOR 1 represented input, weapons input data, and, so, therefore, it seems that the solution, the proposed solution would be to build a weapons data base.

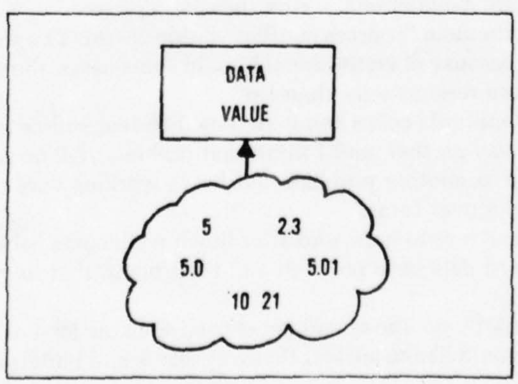
Well, if you'll grant me that that is the solution, for a moment, then you would ask the question, "well, what data would we address in the data base?" Well, the data that would be addressed would certainly have to include at least as the minimum, these models (Slide 28-19). We're looking at models of the future, and given that they have been developed for the most part with the exception of CASM, and that they are on the line, and given that we are going to use them, then we are going to have to have some way of cranking them up. So, as a minimum, we would like to include data requirements that are peculiar to these models.

SLIDE 28-19 MODELS TO BE ADDRESSED IN DATA BASE

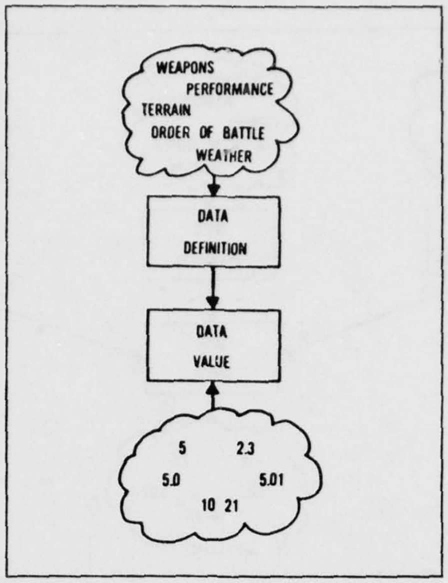
- VECTOR - 1
- VECTOR - 2
- IDAGAM
- TAC WAR
- LULEJIAN
- CEM 4A
- CASM

Well, that doesn't answer the specific question. What would we include in the data base, conceptually? Well, the data to be included in the data base, then, would obviously be the data value (Slide 28-20). That's what we're after. We're after a number. At least we think we're after a number, like 5, 10, 5.01. We need data, but that isn't the whole story. That's the beginning. We have to have data definition. What, in fact, does that number represent? In this case, through a code, we would use a codified system of defining that data, whether it's weapons performance, terrain, order of battle, weather (Slide 28-21). In a weapons data base, of course, we would define it very precisely, very finely.

SLIDE 28-20 DATA TO BE INCLUDED IN DATA BASE

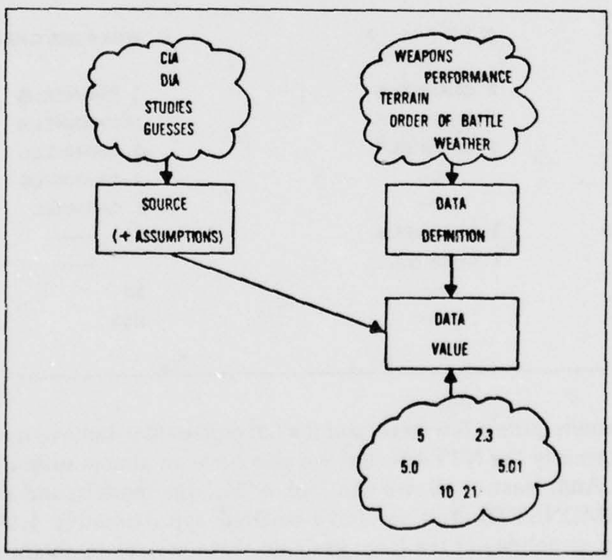


SLIDE 28-21 DATA TO BE INCLUDED IN DATA BASE



Well, we need more than that because this is the answer to your question. There are different types of data and they are based on different sources and different assumptions. We need to codify that and crank that also into the data base (Slide 28-22). Does the data come from CIA, DIA, former studies, or are these analytical guesses? These need to be inputted into the data base.

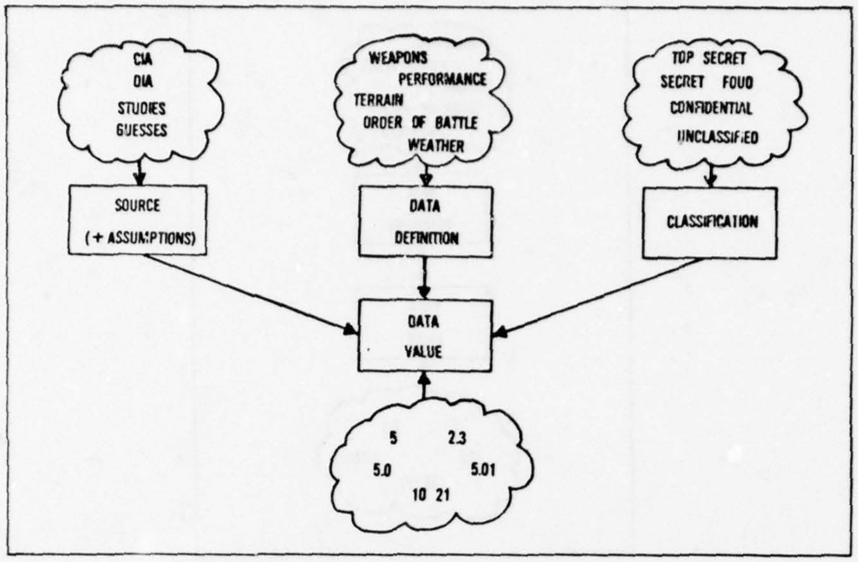
SLIDE 28-22 DATA TO BE INCLUDED IN DATA BASE



There's still another portion that's very necessary, and that obviously is the classification (Slide 28-23). We would have to determine the proper classification of each element of the data base and codify that also. I think now we are approaching something that we'll recognize as being at least an initial manageable approach to storing the required pieces of data. We'll call this a list of data to be included in the data base, conceptually. The sources, the assumptions, data definition, the classification, all link with the data value.

Where are we going to get the data? This is the biggest question of all. Well, we've looked at that problem. We've been looking at it now for over a year, and we have some answers.

SLIDE 28-23 DATA TO BE INCLUDED IN DATA BASE



There are many sources of data (Slide 28-24). This is not the outcome that we've found after looking at the problem for a year and a half. I'll approach that later, but this gives you an idea of the many sources and channels of information that we have.

SLIDE 28-24 DATA SOURCES/CHANNELS

	NTFAM-3	WEAPONS CHARACTERISTICS
AIR FORCE	2 CHANNELS	3 CHANNELS
ARMY	2 CHANNELS	15 CHANNELS
DIA	6 CHANNELS	2 CHANNELS
DNA	—	2 CHANNELS
JMEM	—	1 CHANNEL
MARINE CORPS	2 CHANNELS	—
NAVY	1 CHANNEL	—
MODELS	—	50
DOCUMENTS	—	1100

You notice that we've shown quite a few here, and it's not exhaustive, believe me, for Bob's (Schneider) DoD Force Planning Data Base, formerly the NTFAM, but we also have an almost unlimited amount of channels for the weapons characteristics. And, least of all, we can look at 50 large models, and the data that is produced by these models, such as CARMONETTE, and we have isolated approximately 1,100 documents that contain weapons data. That gives you some idea of the large problem that we have in obtaining data to crank the model.

Well, if we did have a weapons data base, if we did go to the channels, get the information, then hypothetically, what would we want to show as the information content of this weapons data base? I think we want to show these categories (Slide 28-25).

Basically, at this point, we wouldn't want any naval or strategic weapons. Perhaps, we want to add these later on, and we probably want to add things like target acquisition by satellites. But that's looking pretty far down the road.

For each type of weapon, we want to have the primary and secondary armaments, such as machine guns, antitank weapons, etc., and then, of course, we want to show the ammunition, the type of ammunition that each weapon can fire, and actually the physical characteristics, weight, length, fuel capacity, etc.

In the missions area, and this would be primarily for aircraft, we would like to show the mission of that weapon, whether it's NATO low-high-low, or low-low-low profile, etc. Then under production information we would want to list the developer, production capacity, planned weapon replacement, etc.

Slide 28-25 — WCPDB INFORMATION CONTENT

Weapons — Air & Ground Weapons:

- Conventional
- Tactical Nuclear
- Chemical/Biological/Radative

For Each Weapon:

- Primary, Secondary Armaments
- Ammunition
- Physical Characteristics
- Missions
- Production Information
- Target Acquisition
- Vulnerability
- Effectiveness
- Targets

Logical Links:

- Data Classification
- Year & Source of Data
- Weather to Characteristic
- Terrain to Characteristic
- Weapon Ammunition Target
- DoD FPDB

Under target acquisition, we'd show whether it's radar, infrared, range finder, laser, etc., and, under vulnerability, we'd show the vulnerability of the weapon to various rounds or types of ammunition, the mobility, the kills, fire power kills, probabilities, etc.

Under effectiveness, we would show the effectiveness of a given weapon and ammunition, the dispersion, the velocity of a round.

Under targets, we'd show the most likely targets for this weapon.

Now, very importantly, we would have logical links. The logical links would be to data classification, of course, whether it was unclassified, secret, no contractor etc., and then we would show the year and the source of the data, which is very important, and then we would have a logical link, like weather to characteristic. For instance, some data are a function of weather. Other data are a function of terrain. We would like to show this logical link, and then we would like to show the logical connection between the weapon, ammunition, and the target.

And last, but certainly most important, we want to have an automated link between the weapon data base and the DoD Force Planning Data Base, so that in fact you can select a scenario, draw those units and that equipment peculiar to that scenario from the DoD Force Planning Data Base, and then in automated fashion, link to the weapons characteristics and performance data base for the characteristics and performance of the weapons associated with it.

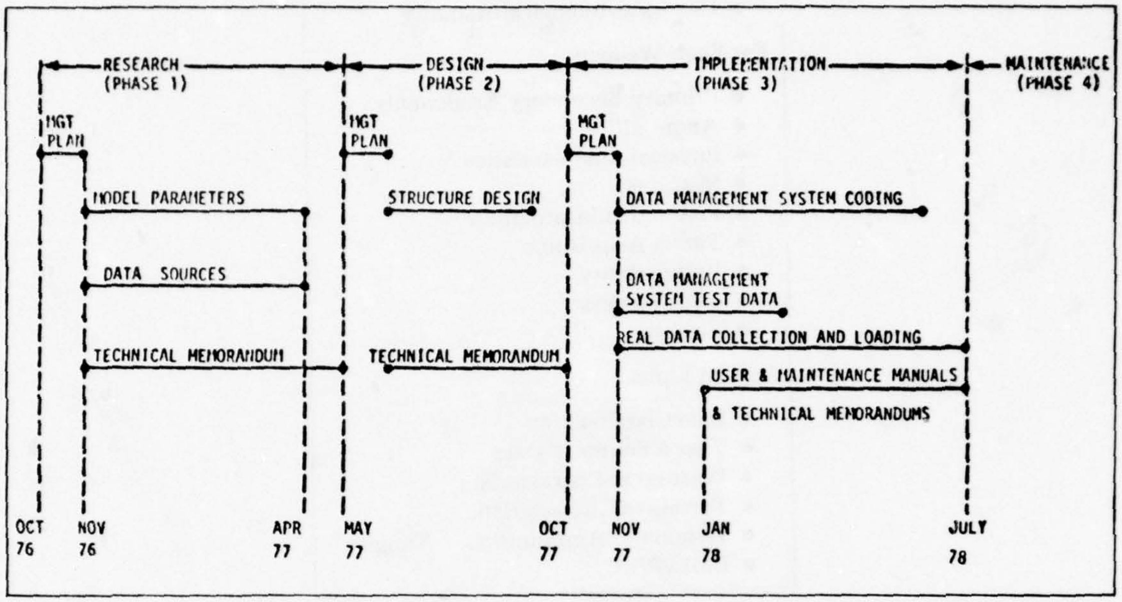
Let's look at a time line (Slide 28-26) and rather than talk hypothetically as I have been, let's talk in a concrete manner for a moment. Is this just talk? Have we spent a year and half in a congressional study, or have we really done something concrete? The answer is that we've done something concrete.

We've gone through a research phase, okay? We have looked at the model parameters. We've decided that there are certain commonalities of weapons data that we need. We have isolated data sources within the Department of Defense, and some outside, and we have produced a technical memorandum. This all started in October of 1976, and by April of 1977, well, in fact, by May of 1977, we had a published version of the technical memorandum listing these sources of weapons data and their availability.

Having gotten by that hurdle, we decided to go into phase two, the design phase, at which time, we talked a great deal about the structure and the design of this data base. After we approached what we considered to be a very viable concept to the design of the data base, we produced a technical memorandum that is at this very moment a living, breathing document showing the design of the weapon data base. I've got a feeling that you'll prob-

ably agree with me that it's quite a viable design. This is something that is workable. Something that we can implement, but not as it is. This document needs to go out to DoD agencies, to people that have the same problem that I do. We would request that you look at it, we would request that you make comments, criticism, and then come back to us with it. We'll revise the design of the data base, hopefully to encompass something that's better, or perhaps something that you would like to see in it, that we have not thought of. Then, we will go into the final stage, which is the implementation stage, and obviously, the most difficult stage.

SLIDE 28-26 SCHEDULE FOR WEAPONS CHARACTERISTICS AND PERFORMANCE DATA BASE



We needed a talking document. Here's a living, breathing, talking document. This is more than anybody else has ever done in the way of data base research and planning, at least on a national level. I think this is a real milestone.

Now, the real laugh in the story here, or course, is that the implementation phase starts in November and ends in July of 1978. Well, that's probably the most optimistic thing that was ever said. No. That part is a lie. We are going to start implementation, we hope, by January 1978, at which time, we will have settled theoretically on at least an initial data base design that's palatable to the Department of Defense, and we will at that point have selected a data management system. We will code the design and we will test the system, and we'll start real data collection and loading. We'll then publish some users and maintenance manuals and technical memoranda which we have to do simply so that we have this system well documented for your use as well as ours.

Well, now that I've let that cat out of the bag, we have a great number of technical problems to talk about (Slide 28-27). Given that we've going to go forth with this, which we are, then we need a choice of a computer. We've solved that already. It's not open to conjecture anymore. It will be the Honeywell 6080, and it will be the WWMCCS Computer. It has great potential for P-winning around the world, and interconnecting, and I think at that point we can see the real enhancement or having a national level data base that can be tapped by others in the command and control chain.

Slide 28-27 — TECHNICAL PROBLEMS — DATA BASE CONSTRUCTION

- Choice of computer
- Choice of data management system
- Structure of data base
- Location and volume of data
- Validation of data
- Connection of low level to high level
- Update of data base
- Security

The cho ce of the data management system is closed. We're going to use the WWDMS, the Worldwide Data Management System, simply because it happens to be a WWMCCS standard and also it happens to be very viable and functional for the weapons data base purposes.

Well, the structure of the data base, I would like to say, is solved, but it's not. I solved it from my point of view in that I can show you the best that I think industry has to offer in the way of a design and I offer it to you as critics to look at, to digest, and help me revise. I think that's fairly settled.

Then there is how to handle the location and volume of data. Well, this is speculation now, but initially what we would like to do is probably throw this thing out in removable disc packs. Then as it gets bigger and bigger, we'll go to tapes, and, eventually, it'll become about like Bob's (Schneider) data base. It will become very large and very unwieldy, but by that time maybe technology will be there to give us a hand we'll be able to do something better with it.

The validation of the data — there's a problem for all of you to wrestle with. I think you are much more capable than I am in this area. What we need are certain edit routines, certain validation routines, that are within the realm of reason, but yet that can give us a reasonable indication of the validity of the data that we're storing, and perhaps point out the gaps.

We need to interconnect, in some technical fashion, the low-level to high-level data. That again, is another area for research. We've done a certain portion of that in the design of the data base — probably, I think, as much as technology would allow. You have ideas on that. We'd certainly be glad to hear them.

Then there is the update of the data base, which is the nightmare of all times, but it's not really that bad. It's manageable because, if you consider weapons, weapons are not constantly changing, except for bringing aboard a small portion each year. We may be throwing aside a small portion, but for the most part, once weapons systems are uncovered, the data remains reasonably stable.

Security is the biggest problem. If there were penetration of a data base, and given that all the weapons that did exist eventually wound up in that data base, this would be a magnificient or monumental security problem. That's something that needs to be dealt with initially, at least, maybe even more emphatically than anything else.

And, of course, naturally, in any type of endeavor like this, we're faced with management problems (Slide 28-28). Who's going to pay for this thing? Certainly not me. Funding is necessary. Well, when you say, certainly not me — every agency says that. Nobody has the big dollars, the millions of dollars, to throw into an endeavor like this to help all you modelers out. About the only feasible solution that I see to it is to pool our resources, and to come to some sort of a mutual agreement as to how the data base will be developed in its evolution, how it will be prioritized and funded.

The data release authorization is a big management problem. Who can release the data from other agencies? Intergovernmental coordination has always been a big problem on the DoD Force Planning Data Base, and it will continue to be on the weapons data base. Naturally, of course, the government versus the private sector is another problem of coordination.

Slide 28-28 — MANAGEMENT PROBLEMS

- Funding
- Data release authorization
- Intra-governmental coordination
- Government vs. private sector coordination

Well, I have here a suggested solution for you to think about (Slide 28-29). I think it would work quite well. I think in an endeavor like this, what is necessary is to set up a Scientific Advisory Group, a SAG. These would be people associated with different agencies in the Department of the Defense that would, in fact, have a great stake in this data base and would, in fact, be tapping into it and using it. Army, Air Force, Navy, sources such as DIA,

Slide 28-29 — MANAGEMENT STRUCTURE

- Scientific Advisory Group (SAG)
  - — General guidance
  - — Problem identification
  - — Prioritization
- Technical Control (DCA/CCTC/C315)
  - — Modeling expertise
  - — Data base expertise
  - — Contract management

CIA, etc., all need to be represented in this Scientific Advisory Group. They would then provide general guidance in the information of this data base. They would identify the immediate problems, and they would provide us with a list or prioritization of how we should approach this endeavor. Right now, DCA is providing the modeling expertise, the data base expertise, and the contract management expertise in this endeavor.

Well, as I promised, I have some recommendations to make in this area (Slide 28-30). I recommend that we link the DoD Force Planning Data Base and the Weapon Characteristics Data Base. It seems only natural that we do so.

Slide 28-30 — RECOMMENDATIONS

- Link to existing data bases (NTFAM-3)
- Incremental approach — i.e., weapons characteristics, followed by terrain data, etc.
- Formation of SAG
- Further study of problems and opportunities
- Communication among gaming community
- Formation of MORS working group

I also recommend that we use an incremental approach and not bite off a multimillion dollar action and then wind up with a disaster. I recommend that we bite off something that we can chew, and that we start in that manner. Perhaps, initially with weapons characteristics, followed by terrain data, followed by weather, etc., etc. — Those things that are being inputted to theater-level models as the state of the art progresses.

I recommend that we do form a Scientific Advisory Group. We have one attempt at this that went off rather well. Frank Kapper initiated the first meeting of a Scientific Advisory Group on the DoD level, and it worked out quite well. At least everybody agreed that there was a problem and at least everybody agreed that there must be some sort of a solution to the problem and there was an underlying commitment to cooperate. The big wrestling problem was that it bounced over to Gene Porter in DP&E to take some action, and Gene left and it left. So right now Bob (Schneider) is here, and I am saying that because the action ball is still there, but I know Bob's up to his ears in other things to do, but that was the first initial meeting. Nobody shot it down. Everybody thought it was a tremendous idea. There was some scepticism. They didn't all get up and cheer, but for the most part, they thought it was a step in the right direction of progress.

I recommend further study of the problem and opportunities, and that's why I'm here. I believe that communication among the gaming community is absolutely necessary. I think we should have open logical debate on this, and if you have a better method, let it pop up and show me, and we'll start pursuing that avenue.

I also recommend that we have a formation of a MORS working group to address this problem. I think this is very significant in the MORS. I think this is one of the least covered topics. Everybody likes to talk about the concepts behind modeling. People would like to drive their car. They like to talk about designing the car, but most people are really turned off at the logistics of having to provide the gasoline and the oil and the checkups.

Well, this is what the data base is. It's more the nastier side of the house.

Let's look at what would happen if we adopted this recommendation and we did seriously form a SAG (Slide 28-31). Well, the first thing we would have to do is to get a charter. The point of the charter is that if you didn't belong to the SAG, and you didn't contribute, and you didn't want to sit in there and help out in this effort, then quite reasonably maybe you shouldn't be able to use this data base. Maybe you should go out and get your own, start key punching the same numbers, spending the same bucks that we're spending. I'm being very cynical on that, but, you know, the point is, if we don't get together we're going to continue this data redundancy problem that we're having.

Slide 28-31 — LONG-RANGE SAG PLAN

- Charter the SAG
- Define data requirements
- Schedule and prioritize acquisition of data
- Coordinate
  - — Funding
  - — Updating
  - — Data release
  - — Future model development

We ought to be able to get together in a SAG and define our data requirements. We all have our own peculiar data requirements, but I think as we found in the first meeting of the SAG, they're not all that peculiar. We should schedule and prioritize the acquisition of data. We shouldn't step into this thing in a multimillion dollar fashion. We should start with something that's realizable. I have an approach there, I think . . . and I have it costed. I have it funded up to and including the design of the data base.

Now, we're talking about implementation. We're talking about big dollars. It takes a lot of money to take this and convert it into machine-readable format and put it out in a data base. I think that the SAG should be responsible for coordinating the funding, the updating, the data release authority, and, as a by-product, I think they'll have a big handle on future model development.

Finally, I submit to you my conclusions (Slide 28-32). In a year and a half of work, which seems like just a short period of time to you I am sure, on the data base problem, I have concluded that the data base will restore emphasis to analysis rather than put it on the collection of the data. This is our biggest problem now. The analyst is out there trying to find the source of the data, trying to get it into machine-readable format, and has very little time for the analysis which is absolutely necessary to be incorporated in this before it's cranked into the model.

Slide 28-32 — CONCLUSIONS

- Data base will restore emphasis to analysis of data rather than collection of data
- Data base will provide greater comparability between models
- Data base will influence further future model development
- Data base will provide identification of data sources and classification for studies
- Data base will identify degreee of data appropriateness to models used in studies

The data will then provide greater comparability between models, and as I promised, Larry, I'm going to say something on validation. The closest I could come to suggesting validation, and nobody has dared talk about it during this conference, is that if we hypothetically had a data base, and if we could crank up several similar models with data from a similar scenario, and made an attempt at explaining the divergence in output from these models, I think that's as close in 1977 as we can approach data validation of theater-level models.

The data base will influence further future model development. If we can show model developers, many of which are present here today, that the data base is the constraint and represents the bounds on the data available in DoD, perhaps we can put a lasso on future model development in generating high resolution models for which data do not exist or may not exist or is too costly to procure.

The data will proivde identification of data sources and classification for studies.

The last point I would like to touch on is that the data base will influence further intelligence gathering by identifying where data is lacking — by identifying the holes that exist. Probably of all these conclusions, I see that as the most significant. If you need the data to crank your model, to do your force planning, and the data don't exist, then I think this will go a great way toward showing those people that procure the data, that we must get the data and we must put it in the data base.

I would like to close by leaving you with the words spoken by Wilbur Payne when he heard about this briefing, he said, "for God's sake, don't talk about it anymore, just do it."

Mr. Schneider: Thank you, Ray. I would ask the various panel participants to come forward, please.

Excuse me, I didn't give you an opportunity for you to ask questions of the speaker.

Maj. Bednarsky: That was very strategic on my part. I thought I'd slip off.

Mr. Steenrod: My question basically runs along the line that Jim Dunnigan talked yesterday about play testing his games. Are you going to play test your data base, and I mean other than just giving the design document to someone to say, "does this look all right to you?"

Maj. Bednarsky: By play test are you thinking of it in an automated fashion that we'll draw from or retrieve from this data base in an automated fashion to input to a model?

Mr. Steenrod: I'm thinking of it in the fashion of someone has a study to run. They have a model they're going to feed and they need certain data elements, that they would, and I recognize that you haven't allowed much time for this to fit in should you decide to do it, but I mean they should set down with that book and see whether or not the data, the logic and the connection of those data items can all be obtained from that data base you have designed.

Maj. Bednarsky: Now, that's what we're hoping it will be; theoretically, that's what should happen. I submit this book, this data design, and people in the industry take it, they look at it, and they see if this logical connection exists, and if not, they make recommendations. But, unfortunately, the reality of the matter is that we're all very

busy. We all have twelve hours of work in the eight hours, and, as a result, I don't expect to have too much of that type of feedback. Any that I get would be of great value because I can't obviously do it all myself.

Mr. Steenrod: Well, at best, you have only allowed two months for that to take place which is a very short time.

Maj. Bednarsky: Well, you know, it's the old case, maybe I draw from the military on this one, unless you present a target date, it never gets done. You can look at some congressional studies. You can look at some FAA and CAB decisions. Okay, I say two months. If I allow twenty months, I don't think I would get much more feedback than I would get in two months. I am being practical now. I think that those people that will look at the data design and seriously attack it technically will call me on the phone within 30 days and say, "hey, I see some problems. Can we have a meeting?" And then instead of papers flying back and forth, we'll go over there, or they'll come over here, and we'll sit down and talk about it, we'll thrash it out, and I think we'll get something going. As you point out, there's good reason to delay. There's no hard and fast rule that says I have to start in January implementing this thing. The point is I have made a target date that seems palatable to the Joint Chief's of Staff, and SAGA, and we're going to try to shoot for that, baring anyone coming up with some problems.

Mr. Steenrod: Well, I maintain that your logical structures will continue to come out for years and years and years, and the thought that you could ever finalize this in two months is very naive.

Maj. Bednarsky: Well, let's say that we would then have something that we would shoot at and we would start off with this. The only demand on the system would be that it be flexible so that we could change this as we went along without great expense.

Frank?

Mr. Kapper: I was just going to make one point Stan, I think your comment is well taken, but by the same token, I think there are two aspects which are very important. I think one is to go back to the management structure that is imposed to make this a viable concept. This was discussed by Ray in some detail. His point, which again, is more specific than your issue, is the fact that I think we really look at this as an iterative process. The point that Ray just made is valid in the sense that the system has got to have sufficient flexibility to permit that, but by the same token, we do have to press on to the business. I think that's the only point that was made.

Roger?

LTC Redelman: The U.S. Army Foreign Science and Technology Center has a parallel effort that is in the design stage also. It will also be on the Honeywell system that will work with DoD. There are a lot of commonalities, but the system that Ray is talking about here, treats not only foreign weaponry but also U.S. weapons.

Mr. Steenrod: Well, my point is only that they're certain obvious logical connections, and all of us can kind of sit down and think of a few right off. But anyone who has tried to draw data from a large data file knows that often the logical connections aren't obvious until you're into a study that requires one. Then, your data base either does or does not need it and it's kind of like the statement from John Bode, once it's aggregated, you aren't going to disaggregate it. So you've always have got that problem facing you.

Mr. Schneider: May we take up that discussion in our panel period because I think that is an important point, and I think, different people are going to have different views about that.

Don't go away, Ray.

Maj. Bednarsky: I almost made it, didn't I?

LTC Doty: Ray, I just wondered, I get the feeling that a lot of the data and the requirements that will be satisfied will be directed primarily, initially, anyway, toward the NATO environment. Has there been consideration for incorporating the Pacific command requirements for data in the data base?

Maj. Bednarsky: Well, okay, a Pacific command requirement. I would assume that if we incorporate national level command requirements, then we would also be incorporating Pacific command, but I think the point you bring up is very well taken. Perhaps, when you go back, you could take a copy of this with you, take a look at it and see what you think, and, if not, get on the autovan and then we'll arrange to come to Hawaii and discuss this.

Anyone else want to have my bulletproof vest?

Mr. Ford: I was wondering if you planned to use this data base in direct support of the gaming? I don't mean that's its entire function, but we have heard that there were four or five different large theater-level games, and that each has its own individual definition of what the value of a unit was because each had been modified in a slightly different way to account for things that hadn't been put into the models. I don't see how a data base can be built in a vacuum. Without thinking about these things I don't see how it can make a large contribution to the computer war gaming effort, although it can certainly provide things like orders of battle and specifics in equipment.

Maj. Bednarsky: Here's where it can make a contribution. When we get to similarities, we get down to raw or low-level data. The thickness of a tank is a thickness of a tank is a thickness of a tank. If we could agree on that, and draw from that, then we're talking about converting this to higher level data. Then I believe your comment is that there are different methods of procedure for aggregating these data and for converting the information to

higher levels.

Well, if we accept that there are certain common data that we're using now everyday that are common and prevalent, and if we could catalog those under assumptions and sources through a code, and identify them in a data base and store it in the data base, then, perhaps, one would agree that maybe for different purposes, one could go in and extract the same data if they agree with your assumptions and your sources. I think that's where it would make a significant contribution. Examples are the MCSSG study, the Central Europe study — things that are all being done by SAGA right now.

In working for SAGA the other problem that we run into is that we see analysts come in, have a three- or four-year tour of duty, and then leave. In the interim, they have created several ad hoc data bases for their particular models, for those particular studies. When they leave, we purge their files because most of the time, we don't understand what these data bases were, where the numbers came from, the sources, the assumptions. This would prevent that waste. We would then be able to input these data with the sources and assumptions into a weapons characteristics data base and we would have that for future use, either to do comparative scenarios later on from the same data, or to be able to look at them, and, say, "look, I want to do a study and I'll basically agree with those sources and assumptions and it would be very easy for me to extract the majority of that data and run my study." That's the new capability which we hope to provide.

Mr. Asbed: I'm going to make a suggestion. I don't know whether you would like it or not, but if you don't believe that you are going to get a response, whether you give them two months or ten months or fifteen months, more than, say, if you only gave them two months. It would seem to me that probably you would not get a response at all. Now, maybe one way to stimulate a response is just to pay for it.

Maj. Bednarsky: By contract?

Mr. Asbed: Look at the possibility that you may have to change your patterns. You know, updating is a very, very difficult and expensive proposition especially for a job of this magnitude. So it might be very cost effective now to pay for it and have contractors look into it.

Maj. Bednarsky: That's an excellent suggestion. You know, there's a constraint on the amount of money we can pay at one time. We can go out and say, "okay, here's 70 contractors for a feasibility study, here's the straw man for the data base design." I assume we'd probably have a fairly good product but there are constraints on this and money, more so than in time. We could do that if we could feel that certain contractors could make a significant contribution after they've read what we've done so far, and, if so, we'd invite those proposals.

Seth Bonder has already made a good proposal, and I feel that he could contribute significantly to this effort. There's no doubt about it. We're looking into that. There are others, too, that have done significant data base work. SPC, for instance; we've been in close contact with Cheryl Herrin and Gary Lucas. They've done studies for Tony Cordesman, so we've used this not in a vaccum, but we've used this as an adjunct to work the contractors have already done for DoD. If there are other contractors who feel they could make a significant improvement, we welcome the effort, provided that we can fund the endeavor.

Panel Discussion — Session V

Mr. Schneider: As I said, I know this is going to be a very controversial subject, and, fortunately, we have some time to discuss it. I am going to ask each of the panelists also to introduce themselves, and I'm going to keep my remarks until last so that I can have the last word.

Maj. McEnany: My name is Major Brian McEnany; I'm from Studies Analysis and Game Agency in JCS, and I'd like to think I fit General Kent's description the other day — I'm one of those guys that do run the models. I was then addressed by some other people who said that field artillerymen can't read slides. Well, that sort of categorized me again, because I am a field artilleryman. I'm also a residual member of what Jerry Bracken called the "gutted SAG analytic effort," and I've also found myself a very small minority member here, and that is as a user of models. That is what Larry Low asked me to come talk about here — from a user's standpoint, some of the uses and perhaps abuses of theater-level modeling, to quote Rex (Goad) in the back. I am a user and I like to think of myself as an informed user because I do understand what goes on inside of that box, both the inputs and outputs, but I have a very small frame of reference. I'm one of the principal guys who worked on the MCSSG study, along with Rex, Klaus Niemeyer, and David Dare, and a host of other people. That conference has been sort of kicked around here and you've drawn, perhaps, some extraneous evaluations and I would just like to set the record straight a little bit.

The MCSSG got started in 1973 when Jim Schlesinger went over and talked to the military committee and the counsel, and he suggested that we ought to do a study on the WARSAW Pact and NATO

force capabilities, and through the international process that got moved to a Military Committee Special Study Group. We went through about five years of work which ran through a data base development, a static comparison, a little bit of qualitative comparisons, and then we finally agreed we'd try to get into some gaming. That's where the UK with David Dare and Klaus Niemeyer, and MOD Bonn, and Mr. Sievert, and SAGA got involved.

It took us two years to come to some sort of a comparative assessment, and it was only through the auspices of STC that we got agreement on a lot of the stuff. But there were several things that did come out of that study, and from the casting of the three models that were used there in the comparison of the five that Bob Farrell put up a little while ago, I'd like to just sort of hold off and say, "all right, a one-day evaluation is fine, but we looked at it a little bit differently, with little bit different measures of capability, and we did it over time." After I get finished, perhaps, you can say, "fine, let's take that and we will make an evaluation and say, IDAGAM, or the RELACS system or Battle Group Model however it comes out." Those were the three models that were used, by the way.

Another interesting thing came out, if you remember General Welch's F-15 Code. You've seen the word IDAGAM on all three days here from various people, and I'd like to point one thing out to you that the IDAGAM that was used by Alan Karr is not the IDAGAM that was used by Lanny Walker to do the four-model comparison, is not the IDAGAM that was used by Bob Farrell to draw his one-day evaluation, is not the IDAGAM that was used by Jerry Bracken in the cost benefit study, I don't think, and was not the IDAGAM that was used in the MCSSG study. So the codes are different; you have the same number but the codes are distinctly different. And I'd like to point out to CAA that I didn't think that Hank Shinol and I had completed IDAGAM II, so if you'd please give us a call, we'd appreciate it.

There were three things that came out of the MCSSG study. If you draw the comparison, the trilateral comparison itself, you're not going to get a terrific amount of methodology because we decided amongst us, that is, among the analysts and the policy people, that we weren't going to put out a lot of methodology. You're not going to get the final reports that show you where the FEBA traces were, or how many divisions were lost. You will see several parameters identified and some results, and that's in a document that I put up on the board there. It's in an international memorandum, that's on the back side, International Staff Memorandum 233-77, and it is classified NATO secret. If you've got the clearance to get into it, it's an interesting document. It has excited attention from the point of view that we thought it was dead. But for the last month and a half I've been briefing that thing to various people up through the decision process, and it's because it addressed a problem that was of interest to alliance members. It addressed things like strategy, and the fact that we were three countries talking about something that was a mutual benefit.

We did go through three phases. We went through a calibration phase, we went through a development phase, and we went through a production phase. But we needed an umpire and that's where Rex Goad and STC came in, and they did a very creditable job. But I said it took two years for us to reach a basis where we could be profitably involved in a comparison between the models.

The three parameters that came out of that, which we think sort of dictated uncertainty in any theater-level model and that Rex pointed out to us, were (1) rates of advance, (2) casualty rates, and (3) the density of forces in combat. Now, of all the things that were looked at in the study, those three appeared to have a disproportionate effect on the results of the games, and I would say that a couple of things are true. The rates of advance were pointed out here, earlier.

We tend to use a sort of fixed body here in the States. Trevor DuPuy started to point to it, and Jerry Bracken came up and said, "yeah, the tenuous basis for these has been pointed out." Leonard Weinstein pointed it out. It's been pointed out by several contractors that rates of advance in theater-level models do not really have a very explicit historical basis, but we use them all the time. There is a question, and that is what is the rate of advance for a theater-level model? Is it to be terrain-dictated where you have the capability of an individual piece of equipment to go across terrain, or is it casualty-dictated that you can only move from point A to B in opposition, or do we need some testing? That's what I'm calling for. I think the testing that's being done at Hunter-Liggett, and with whatever audience influence in TRADOC, that we need testing of maneuver units to capture that sort of information to be placed into a model because I don't think we're using these exercises the way modelers need to capture that information.

I mentioned casualty rates. Casualty rate is just sort of an acronym for bunches of data. We sort of found out a couple of things. Trevor DePuy and others have mentioned the Middle East as a model that you can verify. Jim Dunnigan, I'm sure, has got games that'll come out fairly soon in this area. But, also, if you look at the data that you're using right now, I would ask you where did that data come from, and from what period of time was that data extracted because there are different perceptions as to where the most intense parts of battle took place. If you look at the first couple or three days, you can fire higher rates of advantage if you go to the battle of Kursk and look at the Russian information or the Battle of Goodwood, and look at some of the rates there.

The second thing in the Middle East that came out

was how you capture the data and then some of the judgments that were used, and we got into military judgment. Empirical research and military judgment is an uncertainty, and no matter how you read Clausewitz, how you read Jomini or Napolean's principles, I'm going to interpret it differently because of the situation.

We ran a little survey in SAGA to try to set a scaling constant in IDAGAM, and we sent out about 250 questionnaires to all the senior corps commanders, to the division commanders, brigade commanders in Europe, to the armor school, the war college, the student class, a bunch of analytic agencies in Washington, operations and staff officers, and we got about half of them back that were usable. From that we developed a constant we could use to look at intensity and it was related to losses. Wilbur Payne took exception to the way we did it, but that's neither here nor there. We even got a proposal from Trevor DuPuy to continue the activity.

We did come out with a distribution and we found there was a lot of variability: the corps commander said one thing and the brigade commander said something a little bit different, so everybody's got different impressions as to how things were going to go. But we were able to say that the final results had some sort of a measure of uncertainty, we knew that one parameter there was an uncertainty, we sort of knew what the bounds of that were, and I am not sure we know that in a lot of the data that we're using.

The third thing had to do with the density of forces in combat. That is dependent on the first two, the casualty rates and the rates of advance. But a lot of it has to do with external stuff that comes into the models which has to do with the scenario-related threat data, and in particular, the tactical commitment and how military judgment comes in and commits those forces.

We set it up with the same data, the same threat, the same points of commitment for main attack on D-day, but we left it up to each of us to decide how the game was to be played subsequent to D-day. We all came out with three different answers, but we were in an envelope in which, at the end, we could say, "we're fairly confident that if we move or change the parameter, we're going to move that envelope this way." That's where we ended up in that particular study.

We did identify these three things, primarily through the use of STC. This tactical rate of commitments makes a big difference, and it has to do with the rate the film, as Rex (Goad) was talking about, moves because there is a different perception on how we view it and how other people view the same set of events. That's the different timing that it takes between engagements, between the battle periods. Those are important.

Questions came up between the MOD Bonn and the United States often enough that we had to brief part of the MCSSG to the Army DESOPS last week. The question was, "should I be concerned about it?" And we said, "Yeah, you ought to be concerned about it. The results of these things represent different perceptions from three different viewpoints and you ought to be concerned." We need more investigation in this area, in those three things, the rates of advance, the casualty rates, and density of forces. The fundamental problem is that no matter what model you work in, they're going to stay with you. They're not going away because we happen to use a particular type of game theoretic solution because the three parameters that I've mentioned come in the final answer, and it's the results that we're interested in.

There are some other external factors that come into play. We've been sitting here talking about theater-level modeling, but I haven't really heard anyone speak of what a theater-level model is. We're talking about a model that is primarily ground-oriented. We've had a couple of discussions about game theoretic solutions with penetration and aircraft. Until a model is produced that has an accurate, or I should say, an agreed upon methodology for integration of air and ground, you do not have a theater-level model. I'm not saying that we've discovered a way to do it, because we haven't; we're having terrific problems. But one thing did come out of our study, the joint international study, and that was (and we only examined a little piece of it) we recognize that the full air campaign has got an impact on that game. We only played a little bit of close air support and a little bit of interdiction, and that is what the majority of theater-level models do play now, just a little bit. So you're only playing those assets that are attributable to a corps commander because all we concentrate on is U.S. forces, and it wasn't until three years ago that anyone even said we had to get more detail than an aggregated element for a German division and a French division and a UK division. Now, that information is starting to percolate up, but those are the types of things that the user has got problems with trying to get data. Allied organizations for one. There has been a significant reorganization in both the UK army and the revised reorganizations under Brigade 80 in the FRG and a complete reorganization in the French army. Unless when we jump to the future year data bases we attempt to capture this in some way, we'll have a lot of uncertainty rolling around the outside of those models with the data that we're using now.

We talked about the Europe campaign and I mentioned the integration of air and ground. There have been three large-scale efforts that have been tried within the last three or four years where OSD, the Army, Air Force, and Navy have tried together to get a methodology. Nobody is in agreement because no one has a model that is capturing that data. The last hope right now, I think, is this Project Alpha that is run by

General Dixon and General Starry. That is the group at Nellis that Professor Taylor was talking about and they also have an office down at Langley. You've got the two operators that actually have to work together, have got to come to an agreement, and one of the things that we can do in gaming that they can't do in the services is throw out roles and missions. When you get up and play games with the action officers in the tank, all of a sudden roles and missions become very important and you can't take an aircraft off a carrier and fly it into the ground because the marines get very upset about that and the navy doesn't like to push it around. But, you know, gamers can do that. It's not so true when you get into the real world.

The other point that I want to touch on is doctrine — tactical doctrine. We've got a lot of problems with the new 100-5 that the Army has; it's under fire in Congress and I spent about two hours talking to a guy by the name of Bill Lynn the other day who works for Senator Gary Hart, and he is very much against what the Army has produced. That tactical doctrine translated into modeling means that we have to have the capability of capturing different types of tactics in those models. The models that we have now, particularly sector models or multisector models, have one set of tactics that can be played. But unless you can capture an independent sector and move it and say, "all right, I'm going to completely change tactics for this sector," you're assuming one set of tactics — it may be U.S., or German, or UK tactics, but it is one set. We broached this problem in the study and never solved it. It was understood, but we never got around to saying that we could capture it.

The last point is the data train that went back and forth. I want to emphasize what Ray Bednarsky put out here. SAGA has been funding the design of that data base and again this came out in September 1977 and we do need support. We're continuing to fund this, and I want to emphasize that we need your aid and we need your money if this thing is ever to work. SAGA in the strategic forces has got a concept called the Red Planning Board and that way we try to at least set strategic forces for a series of games. Because we have the mission of putting out a Red integrated strategic operations plan each year, we game that against what's being done for the Air Force and SAC.

We tried to set that up in a study that we are doing now at an advisory level of action officers with the services and we hope that we will have some sort of aid from them in trying to standardize a little bit the matter of what you play. In the standardization of this data base, to answer your question, you're going to have to have a preprocessor for any model that works off of that data base. There is no way you can get around it because you have independent, or model-dependent, factors that have to be squirreled around. The costs as we said are high and we've only got a finite amount of resources that SAGA can put into it. We do need money in order to continue it. But we do think it's a worthwhile project.

I have one last thought and that is about model output. We've had several very interesting descriptions of interactive systems. We've talked about the transparency of the model, but I'm not sure whether you need to make the model transparent or you need to make your analysis sufficiently transparent to the decision maker that he understands what happened. We've found out that one way of doing this is to move more towards the Chinese proverb of "a picture is worth a thousand words." We like to work graphical representations and schematics as much as possible. It makes it easier for us to get across the ideas because we also have not only to look at the results, we've got to come up with some sort of an assessment to make recommendations.

I'd like to close by saying that when we get new guys into SAGA and we start talking about theatre war gaming, the boss comes over and says, "hey, give them a bunch of information." I pick up the documentation and I put it down and I say, "Here's a bunch of stuff that's being done in the community," and I put Dr. Stockfisch's book on the top of that and say "When you're finished with all of this, please read this." Now, I'd like to turn it over to Dr. Stockfisch.

Dr. Stockfisch: I'm Jack Stockfisch. About a year and a half ago I got out of this business after having worried about these problems including particularly the acquisition of data. Part of my credentials, I guess, for being called here is that at one time I was so concerned about the data problem, which was some 15 years ago. I thought a placed called CDEC out at Fort Ord was the wave of the future and then I discovered the hard way that if one seriously did set out to get some data and also to test some hypotheses, it's a very dangerous business. You lose contracts, or you get fired, and things like that and, yet, I still feel it's worthwhile to try to do. However, let me shift the focus now. Enough about my own problems. They're really not my problems anymore. They're yours. However, when I came in this morning to get a feeling for the nature of this meeting, I was reminded, I don't know if any of you have ever seen the wonderful presentation that's been on the educational channels, of "War and Peace" done by English actors, about a 12- or 14-hour series. Everyone should read the novel; it's the best on war that's every been written, I might add. Every once in a while in the story when the generals are talking about their concepts and their plans, which is really the late 18th century version of modeling, the individual in the room who is recounting this, usually it's Prince André, who is the horseholder for General Kutuzov, is sitting there and he kind of goes off in a daydream and he says, "What are they talking about? Does it matter? Does it really make any difference as to how the battle or the war will come out?" Of course, the gist of his arguments is, it won't make a

damn bit of difference one way or the other. And, I must say, several times during this meeting that I've found myself sort of reminiscing like Prince Andre, saying to myself "What are these people talking about? Does it make any difference?"

Now, let's get specific. Suppose a war should break out in NATO in the central front, say a year from now. Will anything that any people in this group have done in following their trade of modeling or researching or whatever it might be have made any difference on the outcome? I think this is a worthwhile question for each of us or anyone in this business to ask themselves. Does it make any difference as far as the human condition is concerned or does it make any difference for that rather poorly educated GI who's in that tank turret, for example, and has to deal say, with enemy tanks that might be coming up and over the hill? I submit that, on the whole, much of this that most of us have been participating in probably won't make any real difference. Had we all been, say, harvesting corn or designing cornpicking machines or looking for oil, I suspect the overall human condition would be somewhat better.

Now, this is a rather sad commentary, but I'd like to elaborate on what I think part of the problem is — I want to focus on data. I'd like to ask what exactly do we mean by the word data. Now, we all have different views on this and I thought I had some conceptions of what the word data meant, but when I heard late this morning that part of that which was being stored in this elaborate process that's being proposed and contemplated were such things as fire-power scores and WEIs and WUVs that really left me rather shocked. I thought people had, hopefully or partially, been educated about the irrelevancies of these particular measures. In other words, I'll suggest that much of this so-called data, as illustrated by the fire-power score, by my conception of the word, is simply not data at all. A number, yes, but by and large, the kindest word I could say for it is that it's garbage and if this is the input for modeling then we are in rather said shape. What one might say is that many of these numbers themselves are the outputs of models, not theater-level models but more fine-grained models, particularly the equations that came out of places like the Ballistics Research Laboratory up at Aberdeen or similar places. In many instances those fine-grained models themselves were not well validated, and all too frequently the numerical inputs that were fed into these particular ballistic equations, which are one ingredient of a fire-power score themselves, the actual physical activity in testing that was undertaken, was often very meager or nonexistent. So, we have numbers, but they are not what I would call numbers in any scientifically validated sense. They may be wrong, or, in some instances, they may be meaningless or both. In some instances, the numbers may be valid, but they may not be a relevant number, and I submit one such number is the concept of a probability of getting a hit. That may be a valid number, but it's not relevant with respect to, say, a measure of weapon effectiveness. Yet all too frequently this is used as one of the key ingredients in many of these campaign models. So I'm suggesting that what we are dealing with when we use the word data is really non-data; it is mainly the outputs of other models most of which have not been validated. Now, I think we are dealing with some rather fundamental problems, and I think basically the problem is due to the immature nature of the discipline that we all are, or, shall I say some of us, have been involved in. This is a most immature discipline. The most immature I've ever encountered. Now, I'm an economist by training and it is acknowledged that economics is an immature discipline in the scientific meaning of the word, compared, say, with physics. But, compared to this batch of stuff, why economics is probably in many respects like 19th century classical physics was when it was in its high lustrous period.

What do I mean by an immature discipline? Well essentially, it's immature because we really have no theory of war. That's the most fundamental difficulty we have, and when you don't have a theory you have no basis on which to deduce models. In other words, you don't have a model until you have a theory. There is almost a one-to-one mapping between a theory and a model and frequently out of that mapping you'll also derive the computational algorithm. But, in this business, there is no validated theory of war so we have no basis for deducing a model from the theory. But that doesn't stop many people who have had, say, a modest training in mathematics or in handling differential equations and what have you. We proceed to model things anyway. By and large that modeling is ad hoc. It's extremely ad hoc. I think, that is the most singular quality of campaign modeling, and, I might add, it's also true for what I call fine-grain modeling where people set out to model, say, a small unit engagement. Now, one aspect of this ad hocery in campaign modeling that is partly abetted by improvements in computational or computer technology is that we go to more and more fine-grained models. So if we don't like the fire-power scores, which were a form of aggregation, we proceed to model things in a more fine-grained way. I think the epitome of going this route is the sort of stuff that Seth Bonder has promulgated, and this leads in turn to the evaluation of models by the realism of their underlying assumptions. We know certain people pursue this approach extensively. Tony Cordesman is probably the hot rock in the evaluation of models as to how realistic their underlying assumptions are, but I might point out that in any respectable discipline, you do not evaluate a model by the realism of the initial assumptions that are employed. Indeed, the purpose of the model is to extract from 98% of the realism those factors which may, in fact, be trivial. But, nevertheless, we do evaluate models by realism

because we have no other way to validate models. The other way, which is what is practiced in most respectable scientific disciplines, is to confront models with evidence, empirical evidence, that is generated independently of the intellectual process of building the model itself. But here is where we run into some serious problems. However, there is a great deal of evidence generated by warfare and there are ways of getting evidence about aspects of combat operation, not necessarily campaigns, but there are ways of getting evidence about specific parts of combat activities — namely, by going out in the field and doing experimentation. Now, I might point out that because our modeling is excessively ad hoc, indeed because we have no underlying theory, this in turn leaves us pretty much at sea with respect to having any guidance on precisely what sort of data should we go out and get. I want to point out here that, in most scientific endeavors when you study their history and their evolution, there is always a sort of an iterative, interactive approach in the formulation of a theory. It may start out as a paradigm and then it gets sharpened into a theoretical formula, into a theoretical form. People then go out and confront the hypothesis that is generated by the model with evidence. From that evidence, people get further insight, which leads them to sharpen and refine the theory. Now, this is a very messy and rather involved process. Then 50 years later somebody neatly writes it all up, puts it in axiomatic form and it all looks very nice, and it's at a point where most everyone, even sophomores in college, can understand it. But, by and large, there is an on-going interaction between theorizing, model building where the model building in turn suggests, well, it's this kind of data we want and this other kind of data we do not need. That leads to the conceptualization of an experiment; frequently, you'll have to wait upon somebody designing instrumentation and what have you. But notice, some guidance is given to the empirical endeavor.

Now, in much of the modeling that has been produced by this community, there is very little that is offered by way of saying well these are the data that we should get rather than that kind of data. Indeed, the answer is, to try and get all the goddamned data you can, whether it's junk or not, throw the information into a computer so it will be immediately accessible to everybody. I think this illustrates something about the overall poverty of this discipline or, let's say in the most generous way, it illustrates the immaturity of the discipline. So there is this extreme interdependence and I've seen very little evidence of people being critically aware of it.

Finally, with respect to data, the data that are going to be gathered and stored at great expense. I indicated in my view that I thought it was either garbage, most of it, or much of it was trivia or irrelevant to what this business is about — namely, hopefully, providing people insight as to how to improve our armed forces should we fight a war.

There is a problem here of quality and I mean quality when that word is applied in some scientific meaning. When we ask about quality of data we should ask such questions as What is the source of the data? Is it from a model? Most of these data are. Now, has that model been validated or has it not? If it's not been validated, then the numbers that are the output of this activity I'd say are not of very high quality. But, we shouldn't be too harsh here; maybe it doesn't need to be too accurate, whatever that means. If the numbers are the result of the model and it has been validated, what was the process of validation? Was it a testing process or and experiment? If so, how was the experiment designed or were these numbers simply the output of somebody's judgment as important elements of the WEIs and WUVs are. As I recall, I forget exactly which of those numbers was which, but I think in one instance the elements that go into the WEIs or WUVs were weighted according to their cost. I think that was considered a favorite number at one time by OSD, and OSD presumably is the highest institution if anything should have something to do with quality control in this business. But on the other hand, if it's not weighted by costs it is weighted by peoples' judgment, although let me submit that military judgment is the most overrated ingredient around. In large part, this is true because you've got literally thousands of individual military officers, unfortunately, exerting their judgment at these levels in the bureaucracy who are usually people who have never been in a cockpit or a foxhole or behind a machine gun or in a tank turret, or if they have it was 20 or 30 years ago. And, there is a tendency as time passes to somewhat romanticize one's own past experiences or to bend his own experiences in a rather self-serving way. Now, I can assert this as true from my own experience — at one time I was shot at on about 30 different occasions in air combat missions. So I know of what I speak, but I must say all the doctrine and recent history that has been cited about the effectiveness of bombing has been egregiously overrated, and I've dropped my share of bombs on enemy targets. But, that is not my business so I can be a little more detached.

Most of the data that seemed to be gathered — they referred to them as weapons charactertistics — I submit is the most important area in which we need data and where our problems are most serious. Most of these numbers, these so-called weapons characteristics, are really essentially physical descriptors or, at best, technical descriptors. Sometimes the assertions of technical performance may be valid provided there were honest engineering tests done, but don't be sure of all the numbers, even those supposed to have been validated by engineering tests resulted from tests honestly conducted or conducted in a detached way. One should ask, Are these measures of performance in some combat setting? And here, again, we have no

adequate numbers by and large because we haven't done rigorous experimentation or evaluation of how weapons perform under actual combat when operated by guys who receive the usual dose of training and are either both scared and fatigued in the context that they're using the weapons. So, a large part of the problem is indeed one of evaluating weapons in terms of their effectiveness. More specifically the problem is getting some insight about whether it is worthwhile to acquire those increments of technical performance in new weapons, increments of performance made available to decision-makers by our research and development community or cartel, or whatever you want to call it. Unfortunately, we don't spend much effort addressing this kind of problem by way of doing testing or serious evaluation because we don't want to raise the awkward question of whether perhaps technological sophistication acquired at great cost may not, in fact, do very much for you in actual combat compared with putting your money into other things, say, like larger numbers of units or more careful and selective crew training and other attributes. But, if you don't raise questions like this, then you can procede to think you can get what you would like to have through technological excellence, whatever that means, and you can elude yourself into thinking that you might be able to overcome his numbers with this other quality.

I would have thought that after our experience in Vietnam we might have learned a lesson on this score, but obviously we've not. If anything, we've essentially refused to seriously consider what Vietnam might have shown us and in turn we indulge in the same myth to an even greater extent as we focus more intensively on the European setting.

That takes care of my unhappy comments.

LTC. Berg: My name is Lt. Col. Don Berg. I'm also with the Studies Analysis and Gaming Agency. I sat down here in hope that I would be able to agree with a number of things that my colleagues up here would have said. Fortunately I was able to agree with one thing that Dr. Stockfisch came up with and that is that the majority of the models need a large amount of weapons characteristics information. But one of the projects that Major Bednarsky has talked about provides us with some idea and some capability of getting some more weapons characteristics information and providing it in a data base. Primarily what we seek to do is to solicit your participation and your assistance in filling out this project and hearing you say whether the project will provide any utility to the community of modeling or whether it will not. The modelers, of course, are way ahead. There are a large number of data bases but they don't all look the same; they're all different; they all have all different things, and everyone knows something about their data base that they refuse to tell other people. We've started a large research project on particular sources, and we solicit your help in looking at that and seeing if you can use it. A lot of effort went into it, and maybe, in fact, some people here can see whether some of the sources of the data that are going to be in the design of this data base will provide them some utility. If they don't, or if you know better sources, then pass that information on to us.

There have been a number of suggestions and comments already as to what good the data base is and what good it isn't. Well, the data base is only going to be as good as the users want and what they seek in their design. The initial setup was essentially a draft proposal. It was designed by technical people, and sometimes the users don't necessarily agree on which way things should be. The user has to look at it and provide that technical person with some additional considerations that maybe haven't been considered. The academic community certainly can assist by looking at those kinds of things that are needed; the methodologies that are needed to take the low-level information and provide high-level information for all the different models. But they also have to document that methodology so that when someone uses some of those high-level data elements he isn't completely blind as to where they came from.

Preprocessors, of course, are needed to take this massive amount of low-level data and process it for the users for the entire spectrum of modeling at all level of aggregation certainly the academic community and the modelers can help in determining the most useful preprocessors. An additional consideration certainly would be that the SAG group be chaired at the DoD level, hopefully by Mr. Schneider, so that, in fact, it can provide some steering, some assistance to the data base design as well as to the user and help them get their project done.

A couple of other things that are going on, of course, that most people are aware of are the DoD study management project, which is attempting to identify all the different kinds of studies and make them known so that there can be an exchange of information and avoidance of duplication. One item in the program has to do with the consideration of data bases and the identification of data bases. Maybe this will help in a project of this magnitude. Hopefully SAG will be able to assist in that area also.

There's one other thing that's going on in the DoD right at this point. I believe that points toward a moratorium on development of new data systems. Now, anything that requires some type of reporting structure has to be readdressed for its cost effectiveness to see if, in fact, there is too much duplication in that particular area. That should assist in development of some of these research projects at this point, too.

Mr. Schneider: Okay, let me just say a few words and take questions. I think Jack certainly gave a beautiful resume of the huge problems in this field. As I said in the beginning our office is a good example of his point. When we were told in 1973 to look into the

data because of MBFR coming on the scene and various problems in trying to understand the really big data problems, we ended up not taking into account hardly any of what Jack has said because there were so many problems so much more demanding in the future than that. For instance, we have a force-plan data base in the DoD. I'm not talking about the WEI-WUVs in that data base because they're not data. They were only inserted because they aren't the data but are just the gist of numbers counts. We have a great deal of difficulty trying to get answers from the various people who build and run models. For example, say we'd like to answer a very basic question about a potential enemy that we really don't know a whole lot about. Well, we're in trouble if we can't get agreement on whether they will take a certain number of days to do something or almost double that. One agency will give us one number, and the other one will give us another number. I'm talking about doubling things, and that can represent a huge force. Then we say, "well, wait a minute, we've got to try to get a handle on some of those very gross factors." So we say, "we're really going to focus on Europe now," which is what 1973 was, a focus on Europe. Since the U.S. force over there is really in the minority that means we do have to take into account NATO and we don't know a whole bunch about all those things. Consequently, there's a huge factor that you have to take into account in the balance between Red and Blue. So, when we talk about these very large problems, we end up focusing on the kinds of solutions that are just as immature and as simple as number counts. But we found that those still are very large factors in seeing how any kind of a war comes out. And, since they were so far off, we never got to these other factors that Jack (Stockfisch) was talking about.

Now, what I'd like to say is that okay, we're in a very sad state of affairs, and I'd like to mention that it's even a little sadder at our particular level at OSD. Really, the name of the game isn't even a question of how the war is going to come out as much as it is a question of what we need for deterrence. That's a whole new area that we have to think about. When we talk about deterrence, we mean those kinds of things that you can see, and so on, but when we're in this sort of situation, the real question is, "how do we go about straightening it out?" Now, I gather, Jack, from what you're saying that all of this stuff that we're doing isn't really worthwhile, and that we need to start over again. But my question is, "where do we start?" How do we start over again, and how do we get the input of a lot of the brains, which have been working on other things, into this particular function rather than relegating it to some OSD office which is more concerned about fighting fires than going off into the future? I might say that I have a very difficult time selling my boss on us doing a lot of this data base work because he says, "it's very important, but that's not our business." He's saying, we've got to worry about other things, and I am sure that's what a lot of people say. So, do you have any answers for that, Jack? Like, "what do we do?"

Dr. Stockfisch: I think 98% of this problem starts in OSD and/or along with Congress, and the problem is more fundamentally one of incentives — indeed, the perverse incentives we have in our military management system. The name of the game is to advocate, to justify getting dollars, and, in our system, it has to be done in a fine-grained way. This starts in Congress, and, if Congress doesn't play this game, OSD plays this game. So, once you intellectualize the process, as it was from 1961 onward, so that modeling and analytical or scientific methodology becomes a language you speak, then you can always control the outputs of the model by controlling the numerical inputs. As a result, the chances of getting an honest number — and that number is one that is researched or a measurement that is researched for a pure and simple motive — are slim indeed. Literally to find out what are the engagement ranges in tank gunnery or to get an honest estimate of the lethal radius of a particular round of ammunition, is close to impossible. For example, I've seen the estimate that comes out of BRL for the 40-mm round; it varies by a factor of 2 or 3, depending on whether, at the time, the Army was advocating an AFIS or something else. The chance of getting honest numbers in this kind of setting where the real war is between the Army, Navy, and Air Force, over dollars, over budgets (and if that war is going on within the Army, it's between the tankers and the artillerymen and the infantry) the chances of getting detached honest analysis, of supporting empirical work in the field, and of getting those field trials done honestly and in a hardnosed way are very slim.

So, you ask what to do about it. The problem is one that people working at our level really cannot do a great deal about because if you try to take it on by yourself, at a low level of organization, you're just going to get your head blown off, organizationally speaking. So, maybe the first order of business is for people to understand more clearly how this incentive system operates, and how it operates in a rather pernicious way, with respect to the broader questions of military management. So maybe it is important to try to get the people at the policymaking level, like your boss, and people up on the Hill, to have some real heart-to-heart talks with themselves because it is they and they alone who are the only ones who can change this basic incentive strucutre. That is as I see it, the sad truth of the thing. I often say to myself I hope to God we don't go to war, because large numbers of systems in our force can only be rationalized through budgetary politics. What happens to our capabilities when we have a force equipped with such systems like the Sheridan Shillelagh or M-60 tanks, which have been equipped with Shillelagh's, or when machine guns

don't work because of crappy ammunition or crappy gun design? If you call that a real serious war-fighting force, I think we're in rather sad shape. The only offsetting redeeming factor is that maybe the other side has a similar set of problems, which they may very well have.

So, if you're going to model things, maybe you should model them in terms of political motivations since different systems are affected by different incentives. I thing that this is the nature of the problem. You ask what to do about it? I say the problem is very basic, very more fundamental. In varying degrees, people in this community have themselves been caught up in this advocacy adversary process, and it's easy to pass the buck.

The modelers, say, "yeah, we know the data's crappy, but it's not our fault." The people who produce the data, say, "yeah, we didn't produce the data for the modeler. We produced it for something else." So the buck gets continually passed, and, also, the people that produce data say, "yes, we're also limited by way of resources." But the numbers producers change those numbers from time to time, and they can do it often, in very good conscience, because there is no ongoing, systematic experimentation and rigorous testing. But we don't want too much testing; we don't want hard information. Because when you get it too hard, the options available to the decisionmakers, which someone spoke about earlier, are not available. If the evidence suggests clearly that you should get this type of airplane, or this type of tank, then you lose some degree of freedom with respect to where the contract goes, whether it's in this country, or some other country, or whether it's in Detroit, or Cleveland, or what have you. So you don't want to be too rigorous, yet you want to provide the assertions that politicians make with some patina of scientific respectibility. Campaign modeling and other kinds of modeling serve very nicely here but that is subjective, I'm sorry to say. Well, that that's the end of this. Question...?

Mr. Schneider: While we're in two different positions, I think we do think very much alike on that. That is one of the reasons why the approach that I've taken in trying to get just the amount of data that we have in the DoD Force Plan Data Base, has primarily been through the camel head-in-the-tent routine. In other words, you have to work with a huge bureaucracy, and one of the ways you work with a huge bureaucracy is to just kind of go sneaking around. A few of those people who are interested in the problem kind of deal with it in a very careful but friendly, informal way, without working through the hard part of the organization. You end up making some headway that way. I think one of the suggestions that I would make if you have an interest in trying to do a job like this is, in spite of the resistance of the bureaucracy and recognizing what Jack has said, to move very slowly but steadily in your area of influence. Typically, while it takes you a lot longer to get there, at least as long as you don't move so fast that somebody recognizes it and bangs you over, you are able to get something done.

For example, within the MCSSG study, because the analysts from the three countries got together a lot in spite of the political problems those three countries had that, in many ways, made them at odds with each other in terms of goals, the analysts were able to talk over the problems. They were able to have some real success in discussing some of the basic problems that were involved simply because it was done on an informal basis. You just kept pushing forward slowly, and while I think that is a very slow way to do things, it may be better than nothing.

Questions?

LTC Redelman: I have a question pertaining to the use of the data base that Dr. Stockfisch, you might say, thought was very poor — the common data base agreed on by the FRG, the UK, and the US in the MCSSG study. I'd like Brian (McEnany) to take a minute to talk about how the human element was put into the use of this data base in using the IDAGAM methodology in the gaming that the United States did.

Maj. McEnany: We talked about the man in the loop. I was going through some notes. All three of those models are run as a man-in-a-loop. It's not automatic — this study was done with a man in a loop. The data base resulted from aggregations of a large-scale automated data base. It was put together as part of what was phase one in the study. Then we sat down around a table, both here in Washington and in London, and talked about the data that came out. We had to make agreements, but we found out that our intelligence agencies, talking about comparable organizations, gave us all three different answers, not so much between the German data, and the UK data, and the U.S. data but all three of us didn't really know what the heck was going on in the Netherlands, or the Belgium area, so that we had to make some agreements. There was some aggregation that was placed on us, and it had to do with effective weapons, which is a problem with these data.

Dr. Bracken: I'm impressed by the comprehensiveness of this approach to the data base development and the data base management system that Ray Bednarsky presented this morning. I would like to come back, though, to the comments that I made the other day. It's true the MCSSG study has been percolating along, and a small, competent group of people from the U.S., Germany, and the UK have been working on it, but if the data base gets developed over the next year, successfully, who is going to work with it? If there are no analysts in SAGA, or very few analysts in SAGA, very few in PA&E, no contract money in OSD, just who is going to use it? And if it doesn't get used next year, or the year after, the system will still be able to accept better data, if better data comes along. So It's

good to have the system, it seems to me, but eventually, even if there are better data derived and developed, who's going to use that data?

Mr. Schneider: There is something that I don't think has been mentioned here yet, and that is that there is a slight ray of hope in all this. It is true that study money is going down, that there is no one really in charge of studies, that it is a very undisciplined area of endeavor, not only in data modeling but in everything to do with theater-level studies, however, right now, there is quite a bit of discussion going on at the Secretary of Defense level about the need for better studies than we've had in the past, even to contribute to the PPBS, Planning, Programming and Budgeting System. I know Russ Murray, my boss, is much more concerned about getting more analysis into the process. That is also true of Secretary Brown, and it's also true of President Carter. I think each of them wants to get more involved at an earlier stage in the decision-making process, rather than at the end, so they can say "I agree," or "I don't agree with what's happening." So I think we are moving back toward, and I say this cautiously, the McNamara approach, but not all the way there. They are very careful to say we're not going all the way back, but that means that there are going to be more things like DPMs where there is some sort of an analytical explanation, at least, of the policies and of what the results of those policies are. It's my estimation that we're moving back more in that direction, so that's a ray of hope. But I can't answer who's going to do all of this, because, as a whole, it is a very undisciplined, and unmanaged community.

Dr. Bracken: Unfunded . . .

Mr. Schneider: Unfunded, and the reason, of course, it's unfunded is because at the top level, I think people have been disillusioned.

Prof. Taylor: I believe I would like to make a few comments. I think, number one, over the long run, we've got to listen to Jack (Stockfisch). I think he has had some influence, he does have reports, and his book "Plowshares into Swords;" I think it all should be serious reading for the intellectual part of our community — read it, digest it, and try to figure out how to operate within the system.

I think, also, that history has many lessons for us. Jack expressed his concern on quality versus quantity. Well, it turns out that there were many conferences on tank warfare after World War II. One document that is now unclassified that makes very fascinating reading is the Fifth BRL Tank Conference from 1953, in which it was especially pointed out that the reason that the German armor was defeated by the U.S. armor was not because of the quality but because of the quantity. We could just keep on pumping out tank after tank — the factories defeated the tanks on the battlefield. Well, so much for that.

Another thing that I think is important for us to critically evaluate. Jack brings up the point, will it make a difference; might not we try to examine, did it ever make a difference? Specifically, there's a very nice myth on the contribution of OR in World War II — operational research, helping along with radar, to win the Battle of Britain. Well, now, it turns out, somewhere around 1974 or 1975 we hear about something called "the ultrasecret." This had to do with a cipher machine that the German high command used for all their coded transmissions to the theater commanders. A Polish guy, who worked in a factory in Czechosolvakia, escaped to Britain, reconstructred the machine from memory, and it was used by the allies against the Germans, meaning that we know damn well where all the attacks were going to be. Operational research was a very good subterfuge to confuse the men. Moreover, Bernard Koopman, who is the father of search theory, recently told us in San Francisco that he was talking to a navy admiral in 1946, and he was told, "well, gee, the best thing that you did was to confuse the enemy on what we were really doing." The *Bismarck* was sunk because, again, the transmissions were intercepted, and deciphered. So I think there is some hard soul searching that has to be done because the myth of military OR, is always, "well, gee, look at how it contributed to World War II." Well, look, I'm a young guy, I got in this business, and I look around me nowadays and I know how well we're doing. Don't tell me that we did any better then, when science wasn't even half as well accepted. The last thing that I want to comment on is that Jack does make recommendations. Maybe he's modest but he has some very articulate recommendations in "*Plowshares into Swords*," that take a while to read and digest, but I think it has got to be on our reading list — that, and his Rand reports that articulate and very carefully document these points.

Dr. Stockfisch: Thanks for the plug.

Prof. Mayberry: I'd just like to make one small comment. I think that one could almost categorize the speakers today, at least superficially, as to whether they were in favor of getting data, and doing war gaming with it, or whether they were opposed to it. I recognize that it is not that simple. Nobody is either in favor or opposed 100%. But I think there has been a suggestion, a statement of fact, which I think is very clear to all of us and that is that every man's judgment may in some cases be biased by the particular circumstances he finds himself in, the expectations of his boss, and just exactly how much of an S.O.B. his boss is, so that he'd tolerate some truth which he didn't happen to find palatable.

In addition to that is a fact that we all have to live with, more or less comfortably, and that is the suggestion that perhaps there is nothing else in a war game, or a study, or an analysis but this sycophancy, and this kowtowing to the expressed desires of superiors. I think that suggestion is really not very appropriate. I'm sure it happens now and then. I'm sure that in this

profession there are a few people who would sell their grandmothers if their bosses ordered them, too, but I think that is not the common situation. I think most of the people in the business are trying as earnestly as they can to make the most honest analysis that they can, even if in a few cases, they have leanings which may be imposed by their circumstances. As someone has said, no man is vertical but there are some men who are upright.

Session Chairmen Summaries

Dr. Kapper: I would like to talk a little bit about something other than just the user's perspective. Although, of course, I will talk primarily about Session One — gaming utility from the user's viewpoint. But, before I do, I'd like to make one comment. I thought Jack (Stockfisch) did a super job in identifying some of the relative strengths, weaknesses, problems, etc. with respect to the data base issue and its relevant problems.

However, I didn't hear him respond with a recommendation for solutions to these problems. I think that is rather key. I would like to say that the presentation by Ray Bednarsky did address what to do about it, and I think as far as we're concerned, with all of its potential and actual shortcomings, it represents to us, Jack, the lighting of a few candles in this whole damn darkness. But I think your points are well taken, and I think we will try to include them.

Now, back to the gaming utility from the user's viewpoints, I don't have anything too profound. Models are tools. That is the way we look at them, and they're just one of several tools in the analyst's repertoire. We need them to do our job because they make our job, as I said before, faster, cheaper, easier, and in a number of cases, a hell of a lot better.

By the same token, they are tools in the conduct of an analysis, or in the conduct of a study. They are input, used, if you will, in the analysis that needs to be performed on a particular problem, or issue. Although those of you who are modelers like to think of the model, and the other things, as the sine qua non for your own personal existence, maybe, the fact remains that we have to use them as tools.

Now, going to back to what has come out in a number of speakers' comments, and, again, I think a point that Jack made is useful: tools ain't either good or bad in themselves; in most cases, it's their use. Models, along with anything else, can be misused for advocacy purposes. If you're the advocate, you're not misusing it. If you're trying to gain an understanding of just of what's actually happening, or if you're trying to go for truth, they can be helpful. But, again, a tool is a tool is a tool, and it's either bad or good, per se, as it's used. It's just like a hand gun, if you will, okay?

Now, I'd like to give you a perspective. I think there are games and models, as seen by the user and as viewed by the modelers, and the beauty, if you will, of that particular object differs because of the beholder. By the same token, there were a number of. charges by users against modelers, you know, "you don't do this, you don't do this, you don't do this," and the modelers, with good justification say, "yeah, and you don't do this, and you don't do this, and you don't do this." There is plenty of responsibility for improvement, and plenty of responsibility for the current problems that we are experiencing.

One of the things that I tried to point out when I led off the session a couple of days ago was to say that we need to talk to each other more and more frequently. I think that is really important. We have problems as users. Our needs change. For example, one comment that was made was, "well, you know, damn it, you said you needed this, and by the time I get it to you, you've got a different need, or the problem is different or something else." That is the reality.

In the context I'm talking about here models take an average of one to four years to develop, if you will, from concept to completion. Martin Shubik, and others, have looked at this over a period of years, and these models cost anywhere from a low of half a million on up. When I say half a million up, I mean way up. But, again, we are tool users, if you will, and our needs change from time to time. One of the things that perhaps you might appreciate is that we really need a range of tools. For some things, we need a damn rough and ready sledge hammer, you know, like QTYP — bang, bang, bang, and away we go.

For other problems, to achieve a particular understanding, we need someting very, very refined. I can name any of the more detailed and very good models that I think that we have here.

Another point I was trying to make that I think is pertinent. Jerry talks about a gutted SAGA. Well, you know, we may or may not be; we're smaller, that's for damn sure. The analytic community overall is smaller. I think, personally, that we are at the low point in terms of our resources, and we're at this point for a damn good reason,

and that is because we haven't sold the value of analysis to the decision-maker, and because we have been victimized, if you will, by Congress, newspapers, and others. We've also been victimized because many of the people in this analytic business, and their bosses — let me put the onus on the bosses — are told, "I want such and such a study, and this is what I want to prove." So the boss is given his marching orders to come up with something that is going to prove that this is bigger, or stronger, or more powerful, or faster, or something else like that. That is a very real thing.

Each and every one of us have had some experience with it, unless you're really pure, and I don't put myself in that crowd because I've been in situations before where the boss, says, "it's like this." And I've been fired at least one time, because I said, "hey, that wasn't what they taught me back there at St. Louis U."

You're going to have to bite the bullet yourselves on some of these things, if you haven't already. In most cases, all you have to say is that Evans Novack would really like to know what the results were or maybe Jack Anderson. But, in reality, I think that, in most cases, you guys have got to stand up and be counted. I'm not now talking to the analysts, but to the users, more than I am to the modelers.

By the same token, it's like truth. When you tell your wife where you were, and maybe you said you're really ticked off, you had a heck of a hard day, so you went over and you had a beer, or two or three, and maybe you had some company as you were having your beer or two. Well, you'll tell her some, but you may not tell her all, and you're not necessarily lying. But you're also trying to avoid a confrontation that may or may not be necessary. It's the same thing when you're dealing with this analytic business of letting the chief know what the heck he wants to know.

Now, irrespective of what some people have said about the military mind and the rest of that stuff, by and large, my own experience with senior military people, with few exceptions — and those guys stand out — they're damn good people. They're honest, and they want to know the truth. If they have to lie, they want to lie on their basis, and on their terms. But they want to get the straight skinny from you as analysts, and you ought to keep that in mind, because that is your responsibility as analysts. You ought not to let people misuse your tools or misquote you.

Well, going back again to the objective of Session One and talk about the types of decisions, etc., where models are used, and also to identify some deficiencies and shortcomings. There is no need, really, to repeat them. You already know those deficiencies as well as I. Models are darn useful. They are a very important part of the process of decision-making, of laying actual plans, maybe not so much in the theater warfare area, but sure as heck in the strategic area. Anyone who is familiar with SIOP knows that. But they're not the sole basis of decision making. There are lots of other things that are input. They do have some shortcomings, obviously, but again, without them, our job would be a heck of a lot tougher, and a lot more expensive.

Dr. Bracken: I'd like to make a few comments to summarize Session Two.

It seemed apparent in the discussion that the state of model development has come along well. The four models are reasonably mature, have been critiqued, have been worried over, and are being used, and are ready for use with the appropriate data. VECTOR II is coming on line. It will require a lot of data. It seems that the data will be developed over the next several years. CASM will probably be coming along in the future. It will attempt to model command and control, and will have a reasonable representation of aggregation and resolution problems. Those developments — VECTOR II and CASM — seem quite reasonable to me, and, I think, were argued to be quite reasonable by the presenters in the discussions.

I would like to make an observation, without any parrot jokes, that, in my view, the logistics modeling, the modeling of combat support, is in abysmal shape. The logisticians are calculating requirements, starting off with the outputs of combat models, not addressing at all the interaction of logistic shortages, logistics policies on the combat. I saw nothing today, and I have never seen anything, that seemed at all mature, shall we say, in this area. But I do think it is a researchable area, and one which should be worked on intensively, as a next step along with the command and control and aggregation and resolution problems.

I worked at the Research Analysis Corporation with Al Tholen, and the Sigmalog people before getting involved with the Advance Research Department, and know a lot of the details of their work. We did the same thing in 1964 as they're doing now, computing so-called logistic requirements. I hate to say requirements without putting "so-called" in front of it. But they were computing logistics support from factors then, and they don't seem to be doing anything differently now.

It seems like a very well organized and serious effort to get the data in which the gentlemen from SAGA and CCTC are getting started, and I'm quite impressed by that, and hope that we make good progress.

It's also obvious that the incentives and study resources need improvment on an absolute basis at the OSD level. As I've said several times, and I'll say again, at the risk of being boring, that there are diminishing resources at the OSD level, the incentives are all wrong. It seems by the economic arguments, or revealed preference that no one wants strong analytical capability at OSD, and I hope with Frank (Kapper), that we're seeing the bottom. But in understanding the incentive structure, I'm quite pessimistic that we have seen it, although, I think the quality

of the work, the little work which is done, is very high, and the MCSSG study is a good case in point.

We had a cut back at IDA in 1972, and it was an opportunity to sort of strengthen the organization because we couldn't afford to keep everyone. So the organization was quite strengthened, on the average, even though it emerged smaller. There's a good side to cutbacks, but if you cut back too far, you're into the bone, which is where we've been lately. We've been letting some quite good people leave the organization, and go off to the civilian sector, and I think we're really hurting.

Next, let me mention the intellectual approach of operations research and defense analysis, to respond slightly to Jack Stockfisch's comments. I was trained, academically, in statistical decision theory and optimization. In graduate school, and later, I became interested in game theory and economics from my colleagues at IDA, and I feel that OR is an excellent way of approaching a defense problems. Military operations, and in particular, two-sided thinking, and the optimization way of thinking, is an excellent language for approaching defense problems.

Jack used the language of operations research in the perjurious sense and with some truth, I think. But it's also a very good paradigm, a very good approach to thinking about OR problems. In particular, at IDA, we've seen a lot of DIA threats that were really dumb, from the Red point of view. By multistage optimization models or any kind of exploration of the strategies available to Blue and Red, we were able to identify game theoretic options which were much better for the Reds, and can have significant affects on the analysis. So I would think that our way of thinking about problems is quite good. In fact, I think it's better in defense than it is in other places because there is a straightforward, two-sided problem. It's us against them. It's NATO against the Warsaw Pact.

In the civilian sector, if you're doing a cost benefit analysis of a highway program, or something like that, and you select one where the benefits exceed the social costs, then distributional problems become very important — who gets the benefits. Take it from this person, and give it to that person, and it quickly goes beyond the capability of straightforward operatons research type of thinking, and, in my view, the economists can't handle it very well either.

Now, though, the physical science paradigm, the approach of the physicists and the scientific method of specifying theories, and getting data, and verifying the theories, that seems to be a litte too difficult in the defense business, too. We can't run wars. The pucker factor is the most important of all, it seems to me. So that, I think, going back and looking at historical data, and attempting to calibrate our models as we calibrate IDAGAM results by attrition and movement curves, based on real people fighting real wars, is a reasonable approach. I think more needs to be done in that area to apply some of our resources. Looking into history is a reasonable way to go. But to actually develop theories and test them and to think that we'll eventually be rich enough to characterize the kinds of weapons systems analysis problems that we have, and the cost effective tradeoffs, is too ambitious, and I'm not hopeful for it.

I think that the language of operations research can go a certain way, but not far enough to be called science.

As a last point, I would like to mention that in my judgment, models of warfare are definitely useful in cost effectiveness analysis, and the development of relative weapons systems, of cost effectiveness decisions, trading off A-10s and F-16s, and tanks, and antitank weapons, and stockpiles of weapons, MAVERICKS — that is something we can do with models. We can do them with multiple models better than with single models. We can do cost analysis, put these things together, and come up with credible results from the point of view of buying systems.

In a way, our job is to help people buy systems within the limitations of the incentive structures, to do as Jim Mayberry says, an honest job with the tools that we have.

Jack (Stockfisch) has been quite critical in the past of modelers selling the models. Maybe that's not a place that DoD should spend money; that's a judgmental issue. I don't know how he feels about it at this point. In the past, I've been quite ready to take DoD money to develop models. I guess I am at this point, too. I do think that command and control, and aggregation, resolution and logistics could be modeled, so, I still think it's worth spending money on. I think I'd rather do cost-effectiveness studies, though, and I think the data development and cost-effectiveness studies are slightly more important right now. I hope that as a community, we're allowed to do cost effectiveness studies, and that there are more than a tiny number of people in the profession over the next five years, although I'm not optimistic.

Any comments, questions, debate?

Mr. Farrell: I'm supposed to be talking about a summary of Session Three, and when I think back to Session Three, I think I can talk about anything at all.

To be serious in that line, Session Three was the only one that really had papers on areas other than simple air and ground combat problems. I think back to both the presentation of naval problems, and, I should say, that Carl Hess's talk, in one of the other sessions, mentioned them at least, and was really addressing them. I think he told us something else, which I'll come to on a different topic. On the presentation on logistics, I agree with Jerry; I think we have paid too little attention to integrating the logistics problem and the rest of our theater-level models.

I'm not as depressed with the state of the art in logisitics, I think, as you are. I think there are models, there,

but I think I have some more information about some of the models that he was talking about, than you do, probably, but I think we really haven't integrated them at all into the overall theater looks we are taking at things.

Even Carl's (Hess) paper, as I understood it, when he was talking about the CAP series of studies, was talking about doing what I would consider a logistic analysis, without a combat analysis. I see Carl nodding, so I think that's correct.

Within the air and ground combat areas, I heard a lot of things in Session Three, and throughout this whole conference, that I think mean we have a great diversity of opinion on some issues. I'd like to talk about a few of those topics, suggest that in some areas it's just a diversity of emphasis and we are, in fact, in agreement, and point out some areas where I think we are, in fact, in serious disagreement.

First, it seems to me, we've had a lot of different suggestions about what it is we want a model to do for us, whether it should replicate reality, whether it should teach us the laws of war, provide us simulated experience to train us, allow us to make optimization analyses of problems, check our reasoning processes against a more formal integration of what is still subjective data, and on and on — provide graphic outputs, a lot of different criteria, against which we would measure the utility of a model. That's where I think, in fact, Carl Hess really showed us what we are all saying which I think is that we don't need one model, and we don't have one utility function for models. We use models to do many different things, even within just the study process, without talking about training, and without really talking about learning the laws of war. In the study process, itself, we need many different kinds of models, and I don't think we have yet solved that problem. We have ad hoc solutions, which we apply, study-by-study, using whatever we can lay our hands on, or create on the back of an envelope. I was quite impressed. I think that was a nice summary of a case study, which I really didn't know that much about, about the way an operations anlayst, in this business, or in nonmilitary OR, really goes about his work. It isn't that he has a nice model constructed by one of those of us who construct a model, and looks at it and says, "ah, that's exactly what I need. I'll put in numbers over here, maybe I'll get them from a good data base that I've had assembled, that has a good management system. Then I'll get my answers out, and I'll look at them, and I will analyze them, and I'll go in and talk to the decision-makers. In fact, when he looks at the model, he says, "you knw, that doesn't exactly apply to what I really want to do in this study." The data base people, I'm sure, can talk about their problems — and did — noting that the data base doesn't always exactly meet study need either. Pretty soon, he has six models and four data bases, and is shuffling very fast. Carl (Hess), at least, gave me some impression of the way that goes, and the way it has gone in my experience, so, regarding what we want a model to do, we have a diversity of opinions. But we are, in fact, in somewhat more agreement that we want it to do many things.

There was some question about what we think our models, in fact, do. It was suggested by some people that we think they really do predict reality, or we act like they predict what would really happen. It's also been suggested that we don't think they give comparative results. I think someone made the point that if you believe that it gives comparisons, you're sort of assuming something about the numbers that are coming out, very much as if you believed it were real. I don't think in that area that I know how to resolve the opinions I've heard expressed. I don't know what we think these models do, and I've nevery heard anybody really give a good statement on that. Maybe, we do have just a multiplicity of opinions and that, in different cases, they're doing different things for us.

An area in which I'm fairly clear that we have very serious differences of opinion is what should be done about validation and the analysis of historical data. I have heard opinions running all the way from, you can't validate against historical data because it's just too difficult and I won't believe it, and it's got a low chance of success, and it's a high cost enterprise, let's not do it to the fact that we absolutely have to do it, if we're to be reasonable about this process, at all. I'm not at all clear that there's any resolution of these opinions; I think we have serious differences in that area.

Along with most of us, I suspect, I have stored up things said by previous speaker that I would like to respond to. I think I'm going to restrain myself a bit, and not talk about data. The data people have had a very good session on data, and will have yet one more chance to summarize that problem. So let me just sum up overall what I think session three and the rest of the conference has said, which is — we do think models do something for us. We want them to do many different things. None of them does any of these as well as we'd like, and we differ with one another seriously about what the relation between models and real historical data is, or should be.

Prof. Lucas: I'm going to be very brief. I think Martin Shubik and Jim Mayberrypointed out very clearly some of the problems in game theory, some of the problems using game theory in modeling and some of the misuses of game theory or places where they could be improved.

I think it came across very clearly in our session, that a lot of the analysts think there's a tremendous amount of theory out there that has not been tapped. I guess the suggestion is that if you haven't really looked carefully lately, in the right places, you probably should try to do so. Some of us feel that communication has broken down quite a bit. In fact, that's probably the biggest problem. The clearest insight I get out of the meeting is there's this great, great need for communication. It's clear that there is a need for more communication within the gaming community itself, but it's particularly true that we must increase communication between the gaming people and

the analysts, game theorists, or other types of analysts. I think game theorists have been lax in not communicating to potential users some of their new discovereis and where they may be used. On the other hand, a lot of these new discovereis are by very young people who are probably not aware that there are these potentials and possibilities. So somehow just opening up these possibilities may lead to some better results.

Conversely, I think the gaming community has perhaps kept the same old consultants around too long. I don't want to use that in any negative sense, but you can start asking questions about how many names have been added to your consultant list lately, how many of your consultants are under 45, or some other types of questions. I've done this with some of my students and some faculty members at Cornell who have gone out to industrial governmental situations. I said "tell me how many young consultants they have or new consultants?" I don't mean young, necessarily, but new to the organization. Several have reported back, and the names are those that followed World War II, or those first few years after.

I'm not complaining if your organization has George Dantzig or Tukey, or someone like that as a consultant. I'm not suggesting you drop them, but I'm concerned with some additions, on the other hand. I think this is a problem and that we have to work from both ends to try to open up communications.

I think a lot of the value of gaming and modeling is educational, learning. This is kind of accumulative in nature, and a lot of little trivial facts eventually add up to quite a body of knowledge. I do think from gaming exercise, we probably learned a lot of things that we acknowledge are negative. We may have found a lot of bad strategies and we've eliminated them from some of our considerations. That is still learning, and it is worthwhile information; we shouldn't forget that.

Often the decision-maker can use the modeler as an excuse for backing up the position he wants to take anyway, and, if you feel you're being used that way, you just have to stand up against it. I certainly know a lot of young people who have been successful in their careers, and I don't think they would let themselves be used in this way. They're aggressive and honest, and would stand up for what is right in that. But if a particular area of modeling is getting too much of this type of thing, then somehow it's a very serious moral problem, and I hope that people feel secure enough to face up to this and respond. Again, maybe, there has got to be some young people who don't have quite as much on the line when they stand up to the boss, as do most of us who get good salaries, and maybe have children in college or something like that. But it can be a very serious moral issue, and it's got to be confronted.

I was a little disturbed about hearing that numbers are garbage, and that we have no influence on what's really going to happen, and that we're a very immature discipline that compares so unfavorably with economics or physics. Hieronymus tells us that there is nothing new, except that which has been forgotten. Well, that's bunk. In mathematics, the literature doubles every eight years. Do you know any theorems that are less than sixteen years old? That's two half lifes back. We are gaining knowledge, and we shouldn't put these things down the way we sometimes do, although we should have good arguments and debates.

These are most of the points I wanted to make. Again, I think we had a couple of very good speakers who indicated the importance of the analysts interacting with the modeler. I mean, the analysts should interact at different stages. Every once in a while, you talk to a statistician, and they say, "they gave me this data, and now they want me to analyze it. Why the heck didn't they call me in when they designed how they were going to take the data? Then I could have gotten some data I could have really used." So similarly the analyst should get in at the early stages of the model-building process. He should come in at later stages, too, and even in the modeling and practice. I believe strongly in the man interacting with the model; we had a couple papers on that. I think some of the most successful events will be when the decisionmaker or someone can interact, and the human factor can be involved, along with the computer. But the analyst should come back at different stages in the beginning, towards the end, and see if he can't help out in these particular types of problems.

Mr. Schneider: Since everyone else has spoken of data, I guess I don't need to, but I will anyway.

To summarize, I think a major point that has been made is that model development is far ahead of data development. Therefore, if we are going to continue to do something about this whole business of modeling, operations research, systems analysis, and so forth, we need to focus on the development of data. I think that means that there is another problem, and that is, that the data gathering we do right now, is focused in the wrong direction. It's not necessarily focused in the wrong direction for the data gatherers because they're doing what comes easiest, but in terms of being able to use that data. As I mentioned, our intelligence gathering is just not focusing our resources in developing the kind of data that we really need to collect for our models. That means that we either have to change the models, or we have to change the data gathering business.

The third point, I'd like to make is that, within the modeling community, we have to learn how to live with the less than satisfactory situation we've been discussing and to do something about it, because I think we've pretty much agreed that at the fairly high level, decisionmakers are doing things which they think are best. They're not going to change just becuase we tell them to, and the way that I have found to be successful within any large bureaucracy, recognizing that you've got many different proponents, is that enough people who have that concern

about a particular issue need, by themselves, first to recognize the potential of getting something done, and then set about to doing it.

I believe that those of us who are really concerned about doing something can do it, but it has to be done on an informal basis. I don't believe we're going to change the formal system. I believe Jack (Stockfisch) is right there. We're not going to change the formal system, but that doesn't mean we cannot do something about the informal system. Someone mentioned here this afternoon that DoD is one of the last great entrepreneurs around because a person within DoD can get a very large project done that costs much money simply because of his tenacity and patience in keeping something going. This is a very good approach that the Communists have developed — the pushing door. You have to keep the pressure up, you don't go too fast because you might get knocked down, but you just keep pushing until you get something done. That is done through the informal part of the bureaucracy, not through the formal part. So what happens, through this kind of patience, prudence, diplomacy, is that something gets done. It means that there has to be an informal communication going on among the people that are wanting to do the job, and it means that at times you need to get together to form a lobby, if necessary, to move out.

I think one of the areas that SAGA and CCTC is working on right now is an example of just that, in the development of this data base, in spite of the fact that at a fairly high level, people are not that much interested. I still don't know if that is the proper way to do it, but I certainly believe that we should be working together to get the job done.

Now, I want to inject one new idea that I just thought about. It may be that because this is such a huge problem that, perhaps, and this might be done through the SAG or in the same way that this conference came about, we should have a data symposium, something in which various data providers and data users actually get together and talk about this problem.

For instance, I noticed quite a lack of many of the data providers, who have all kinds of problems they'd like to talk about, the DIA people, for instance, the analysts, and so on. That might be something that we should think about, and, perhaps, we can talk about that in disucssing the SAG.

It does appear as if the data problem is the biggest problem that has been identified in the conference, and we need to do something about it. But I'm convienced that the only way we're going to do something about it is if we informally decide it's important enough; that regardless of where we get our money, we're concerned enough about it to do something because it's our business.

Mr. *Low:* Well, the unbelievable has happened. We've finished this meeting, and it's four o'clock, and I just have one thing to add to all the things that have been said here today, the last three days, as a matter of fact. When I started out with the notion of this meeting, it was the concept of getting this diverse group together that kind of intrigued me. But I must admit that I had many trepidations about just what results of the interaction would be. I must say, that I found it to be a very interesting, enjoyable meeting, and, as I have told several of you, I would hate to think that it was created just for my own pleasure and edification. I only hope some of you share my feelings with respect to the meeting and its content.

As Bob Schneider has indicated, serious thought should be given perhaps to, splinter groups developing from this, and, certainly, keeping up the communication, though, between the various groups. In that regard, several people have asked me to compile a list of the attendees, their addresses, and phone numbers, and see that all of you get this. I'm sure you've had an opportunity to meet a lot of the people here, and might want to continue the relationships that might have started here.

To close, I want to personally thank each and very one of you for coming, and those of you that leave the Washington scene, I hope you have pleasant journeys home, and thanks very, very much. The meeting is adjourned.

APPENDIX A

DYGAM: AN ALGORITHM FOR SOLVING MULTI-STAGE GAMES

by
Zachary F. Lansdowne

28 September 1977

CONTROL ANALYSIS CORPORATION
800 Welch Road
Palo Alto, California 94304

DYGAM: An Algorithm for Solving

Multi-Stage Games

by

Zachary F. Lansdowne

This paper describes a computer programming system called DYGAM (DYnamic GAMes solver) which is designed to aid in the solution of a broad class of multi-stage decision problems including multi-stage games. This system is written in FORTRAN, is based upon dynamic programming, uses polynomials to approximate the payoff surfaces, and then uses a forward "look ahead" algorithm to improve upon the accuracy of these surfaces. Numerical results are given for two multi-stage game applications.

1. INTRODUCTION

DYGAM (DYnamic GAMes solver) is an experimental program written in FORTRAN which can be used to solve a broad class of multi-stage decision problems, including single person deterministic problems, single person stochastic problems, and two person games. The basic documentation for DYGAM is given in Reference 1. The program is modularized into two sets of routines: a master set and a user set. The set of master routines remains undisturbed when changing from one problem type to another, whereas a new set of user routines must be developed for each application. Some references which have compared DYGAM with other solution algorithms are: Bracken [2], Fish [3], Dresher [4], and Anderson et al. [5].

This paper gives a brief description of the approach used in DYGAM for solving multi-stage games. We are concerned with games played between two sides, referred to as the Blue and Red forces. These forces could refer to ground troops, as in the two-piston game formulated in Reference 1; or they could refer to air forces, as in the Berkovitz-Dresher game model in Reference 6. A game is played over a fixed number of periods or stages. Each period could refer to either a day, half-day, etc. We will consider both sequential and simultaneous games. In a simultaneous game, both sides redeploy during each stage; and in a sequential game, one side redeploys during a stage and then the other side redeploys during the next stage. The payoff to each side at the end of the game is a function of the numbers and locations of the remaining forces. The problem faced by each side is how to deploy and manipulate its forces in order to maximize its payoff at the end of the game. The payoff to Red is assumed to be the negative of the payoff to Blue, so that we have a zero-sum game.

Sections 2 - 4 discuss our approach for solving multi-stage games of this kind, and it is based on dynamic programming. The main conceptual limitation of dynamic programming is the dimension of the state space. In order to solve a multi-stage game with dynamic programming, it is necessary to determine (or estimate) the optimal payoff for each state and stage of the game. So as to enlarge the class of problems that can be solved, our approach is to approximate the payoff functions with polynomials; this means that it is only necessary to explicitly determine the polynomial coefficients for each period, rather than the actual payoff at every feasible combination of state variables. However, this approximation approach may give inaccurate answers due to errors in the approximation. Hence Section 3 suggests that the polynomials should not be used directly to provide the payoff estimate for each initial state; rather, they should be used as guideposts for a forward evaluation. In general, one would use the early polynomials to eliminate most of the potential strategies at each stage, in order to allow a search-in-depth on the remaining strategies using the later polynomials. It is also desirable to estimate the error build-up in the polynomial surfaces, and a method for doing this is given in Section 4.

2. DYNAMIC PROGRAMMING

Let $S = (S_i)$ represent a N-dimensional vector of state variables which completely characterizes the status of the game at the beginning of any stage. Refer to Reference 1 for examples of state variables for certain multi-stage games. We assume that there exists a known payoff function $p^{(T-1)}(S)$ which gives the reward to the Blue army if the game ends (at the end of the T^{th} stage) in state S. We also assume that the payoff to the Red army is $-p^{(T+1)}(S)$, so that we have a zero-sum game. Define $p^{(1)}(S)$ to be the optimal payoff (if both sides make their best moves) to Blue if the initial state is equal to S at the beginning of the game. Thus $-p^{(1)}(S)$ will be the optimal payoff to Red. Our problem is to compute $p^{(1)}(S)$ for any initial deployment S, and in this

way evaluate alternative initial deployments.

Rather than directly computing the optimal payoff $p^{(1)}(S)$, the dynamic programming approach* first computes the intermediate payoffs $p^{(t)}(S)$ for $t = T, T - 1, T - 2$, etc. Here $p^{(t)}(S)$ has the interpretation of being the optimal payoff to Blue when starting from state S at the beginning of the t^{th} period. The method begins by determining the best strategies during the last or T^{th} stage. For example, suppose that the state of the game was S at the beginning of the T^{th} stage. Then only a one-stage problem need to be solved to determine the value $p^{(T)}(S)$ of the game. If a simultaneous game is being played (both sides redeploy during each stage), then a single stage game must be solved, perhaps by linear programming. But if a sequential game is being played (only one side moved during each stage), then the optimization problem is either a simple minimization or maximization. Unfortunately, it will be necessary in general to determine or estimate $p^{(T)}(S)$ for all possible states S. The next step is to determine the strategies during the second to last stage (stage $T - 1$). Again, only a one-stage problem need to be solved to determine the value $p^{(T-1)}(S)$; here, $p^{(T)}(\cdot)$ is used to compute the relative value of states at the beginning of the T^{th} period. This process is repeated until the first stage is reached, where the function $p^{(t)}(\cdot)$ is computed from the function $p^{(t+1)}(\cdot)$ for $t = 1, \ldots, T$ by solving a single stage optimization problem (a single stage game for simultaneous games, and a maximization or minimization for sequential games). Note that the optimization proceeds backwards in time, starting from $p^{(T+1)}(\cdot)$ and ending with $p^{(1)}(\cdot)$.

In dynamic programming it is generally necessary to determine (or estimate) $p^{(t)}(S)$, $t = 1, \ldots, T$, for all possible states S. Rather than determining these payoff functions exactly, our approach is to use a polynomial with sufficient degree to represent this function. This device enables us to enlarge the class

*For a general description of dynamic programming, refer to either Bellman [7] or Bellman and Dreyfus [8].

of problems solvable with dynamic programming. Basically, there are two approaches for fitting a polynomial to a function. One is to use a single polynomial with high degree to approximate the payoff function over the entire state space; this approach was investigated in Reference 9. The second method is to subdivide the original space into smaller regions and then use a separate polynomial of lower order to approximate the payoff function over each region; this latter apprpach is called subregional approximation and is used by DYGAM. We will employ the notation that $p^{(t)}(\cdot)$ refers to the optimal payoff surface, while $Q^{(t)}(\cdot)$ refers to the approximate surface. Refer to Reference 1 for a more complete description of our polynomial and subregional approximation technique.

Rather than using the original state variables $S = (S_i)$, it may be desirable to write the polynomial expansion in terms of some function of S. Furthermore, it may be desirable to allow this transformation to vary from stage to stage. Let $Z = F_t(S)$ be a vector-valued function which transforms the original state variables at the beginning of the t^{th} stage into the new variables Z. It is not necessary for the vectors Z and S to have the same dimension. For example, one approach towards solving a problem with a large number of state variables S would be to define new variables Z such that Z "approximately" characterizes the status of the system but with a smaller dimension. Thus, instead of approximating $p^{(t)}(S)$ with a polynomial written as a function of S, we will allow it to be approximated with a polynomial written as a function of $F_t(S)$; i.e., our task will be to choose coefficients for the polynomial $Q^{(t)}(\cdot)$ so that $Q^{(t)}(F_t(S))$ approximates $p^{(t)}(S)$.

In summary, DYGAM enlarges the class of problems solvable with dynamic programming by incorporating the following features:

a) The optimal payoff function for each stage is represented by a continuous approximating function;

b) A subregional technique is used in which separate polynomials are fitted for separate regions of the state space, rather than

having one polynomial defined over the entire region;

c) The polynomial expression can be written in terms of a transformation of state variables, rather than the original state variables.

The user of DYGAM is given wide latitude in selecting the approximation and transformation functions to be used.

3. FORWARD EVALUATION FOR SEQUENTIAL GAMES

The polynomial approximation technique may give inaccurate answers due to errors in the approximation. What is proposed here is to use these polynomials with forward evaluation. Reference 1 shows that this combined procedure will provide accurate answers even with inaccurate polynomials. Our approach will differ for sequential and simultaneous games, as the first case involves pure strategies while the second case involves mixed strategies. In general one would use the early polynomials to eliminate most of the potential strategies at each stage, in order to allow a search-in-depth on the remaining strategies using the later polynomials. The method of forward evaluation used in DYGAM for sequential games is discussed in this section, while the method for simultaneous games is omitted.

In a sequential game, each side moves alternatively. We assume that the approximate payoff surfaces $Q^{(t)}(\cdot)$ have already been constructed, where t refers to the stage. The forward algorithm in the sequential case is characterized by two parameters:

$$w = \text{the width parameter}$$

and

$$d = \text{the depth parameter.}$$

The outputs of the forward evaluation are: the best estimates of the optimal game payoff for a given initial state; and a sequence of moves for each side

which will yield this payoff.

The general idea is to look d stages forward along w paths from an initial state. The payoffs at the end of these w paths are compared and the best one is retained. This best path is traced back to the initial state to determine the first move in the forward direction, and the whole process is then repeated. These steps are illustrated in Figure 1 for w = 5 and d = 4. Suppose it is Blue's move during the first stage. Let S_1^1 be the initial state; and let S_k^2, k = 1, ..., w, represent those states at the end of the first stage corresponding to the w strategies with the largest payoffs, when evaluated with the polynomial $Q^{(2)}(\cdot)$. Next, examine all possible Red moves from each of the states S_k^2. Let S_k^3 be the state at the end of the second stage corresponding to the strategy with the smallest payoff (since it is Red's move), where this strategy is applied to state S_k^2. Similarly, let S_k^4 be that state at the end of the third stage corresponding to the strategy with the largest payoff (since it is Blue's move), where this strategy is applied to state S_k^3. And let S_k^5 be that state at the end of the fourth stage corresponding to the strategy with the smallest payoff (since it is Red's move), where the strategy is applied to state S_k^4. Because in this example d = 4, the forward evaluation has completed its search d stages deep. The next step is to compare the relative magnitudes of the polynomial payoffs $Q^{(5)}(S_k^5)$. Since it was Blue's turn to move during the first stage, the best path corresponds to the one which maximizes $Q^{(5)}(S_k^5)$. In this illustration, suppose that the best path corresponds to k = 2; then the strategy taking S_1^1 to S_2^2 is retained, and the whole process is repeated starting from state S_2^2 at the beginning of the second stage, as if S_2^2 were the initial state. This is illustrated in Fig. 2. Here the w best strategies at state S_2^2 are retained, giving rise to the S_k^3 for k = 1, ..., w. Then the algorithm goes forward d stages deep, giving rise to states S_k^6 at the end of the fifth stage. Since it was Red's turn during the second stage, the best path corresponds to minimizing the payoffs $Q^{(6)}(S_k^6)$.

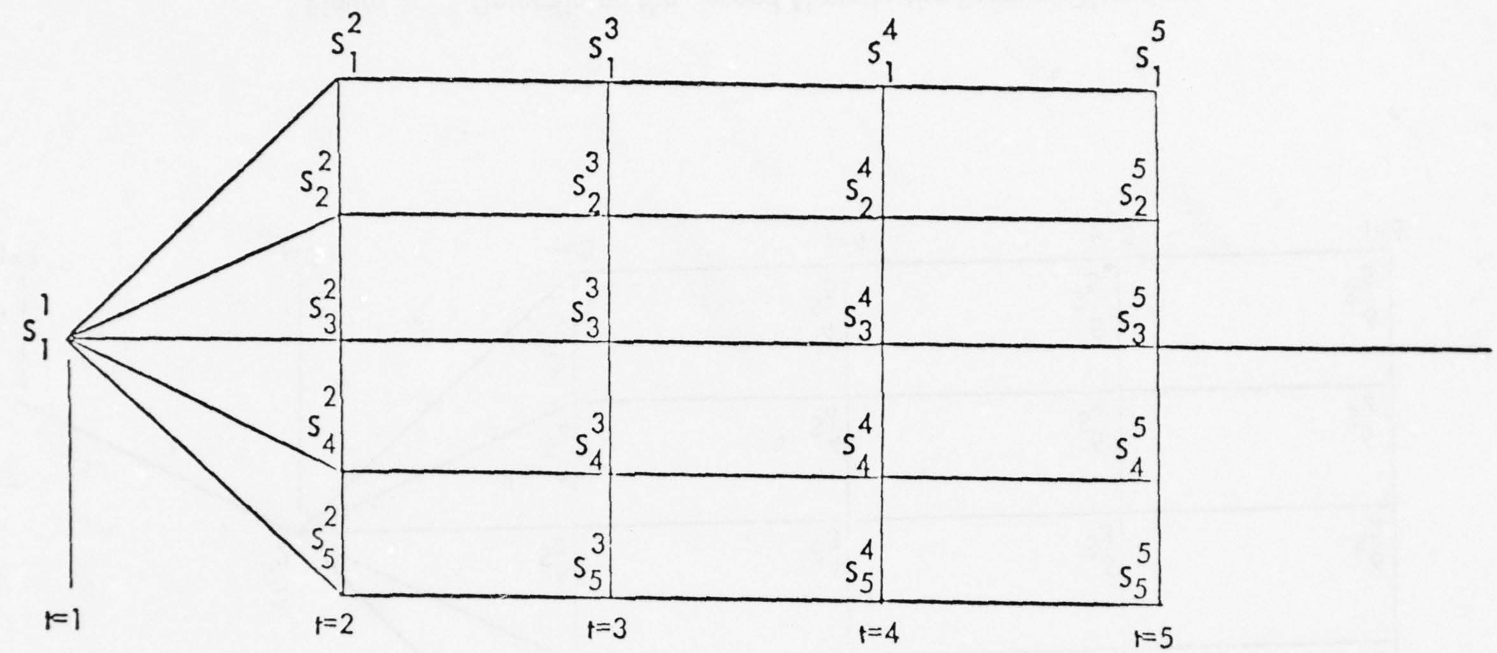


Figure 1. Determining the First Move In the Forward Direction

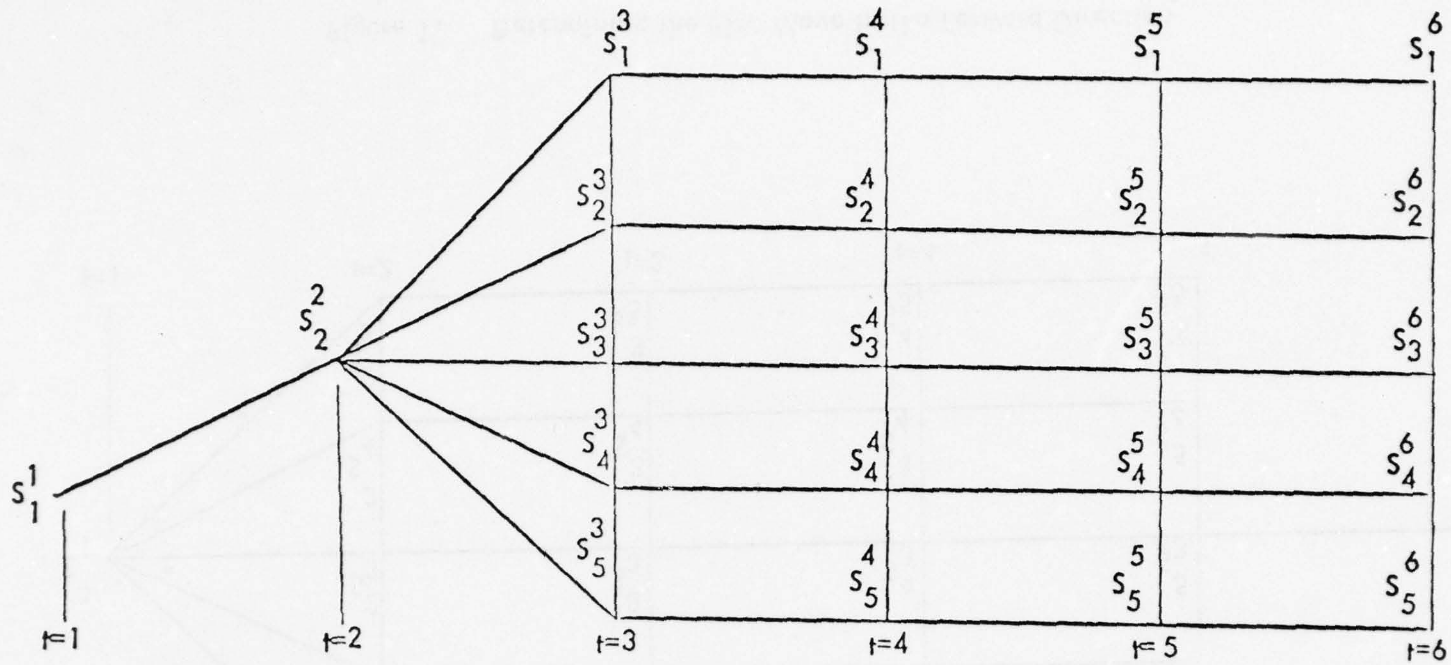


Figure 2. Determining the Second Move In the Forward Direction

Suppose that k = 3 provides the minimum. Then the strategy which transforms the state from S_2^2 at t = 2 to S_3^3 at t = 3 is retained, and the whole process is repeated starting from state S_3^3. This is continued until the last stage of the game. In this way, a sequence of moves is traced out starting from the initial state S_1^1 and ending at the last stage of the game. The actual payoff for this final state becomes the payoff estimate for the initial state S_1^1.

DYGAM allows the user to select arbitrary values of w and d when using this algorithm. Reference 1 includes the description of the forward algorithm in the simultaneous game case.

4. ERROR ANALYSIS

As discussed in the previous section, the forward evaluation for sequential games searches d stages forward on w paths in order to make a single move, and this single move is made by comparing the rleative magnitudes of the payoffs d stages deep (i.e., choosing either the largest or smallest payoff). However, it may be possible to say that a given path has little probability of being the best one, without going the full d stages on this path. In order to do this, it is necessary to estimate the amount of error that arises in the polynomial surfaces from one stage to another. Our approach to this problem is to represent the accumulation of error with a random walk model, as discussed in this section.

Suppose that the forward evaluation is being used to determine the best move at stage t. As discussed in Section 3 our approach is to search d stages forward on w paths. Let

S_k^v = the state at time v for the k^{th} path during forward evaluation ($1 \le k \le w$ and $t+1 \le v \le t+d$);

$X_k^v = Q^{(v)}(S_k^v)$ = the value of the polynomial at stage v for the k^{th} path during forward evaluation.

Thus the forward evaluation computes X_k^v for v = t+1, ..., t+d in order to select a move at stage t. If the polynomial surfaces were perfect, then X_k^{v+1} should be

equal to X_k^v . However, these surfaces are only approximate, and so there will be an error ε_k^v, where

(1) $$X_k^{v+1} = X_k^v + \varepsilon_k^v .$$

Our approach is to estimate the mean and variance of ε_k^v by assuming that the distribution for ε_k^v is independent of k and v. In other words, our assumption is that this error build-up will be the same between any two adjacent surfaces and for any areas on these surfaces. During forward evaluation, new values of X_k^v are being observed for increasing values of v and t. Using these data, the mean and variance of ε_k^v can be estimated with standard statistical techniques, and these estimates should be updated as new data becomes available. Define

μ = the best estimate for the mean of ε_k^v,

σ = the best estimate for the standard deviation of ε_k^v.

Suppose that the forward evaluation is determining the best move at stage t and that it is Blue's turn to move during that stage. Suppose further that $X_1^v \geq X_k^v$ for all paths k and for a specific value of v. Since we eventually wish to select that path with the maximum payoff at stage t+d, we would only wish to cut the k^{th} path off at time v if we could be assured that $X_1^{d+t} \geq X_k^{d+t}$. Thus we have the following problem: given the values of X_1^v and X_k^v at time v, what is the probability that $X_1^{t+d} \geq X_k^{t+d}$ at time t+d? For convenience, we will assume that ε_k^v is statistically independent of $\varepsilon_{k'}^{v'}$, for $v \neq v'$ and $k \neq k'$. We will also assume that

$$\varepsilon_k^v + \varepsilon_k^{v+1} + \ldots + \varepsilon_k^{t+d-1}$$

is normally distributed with mean $(t+d-v)\cdot\mu$ and variance $(t+d-v)\cdot\sigma^2$.*

---

*This assumption would be justified by the Central Limit Theorem, if (t+d-v) were large.

The foregoing assumptions have not been verified with empirical data. These assumptions imply that the conditional probability that $X_1^{d+t} \geq X_k^{d+t}$ (given the values X_1^v and X_k^v) is given by

$$\Phi\left(\frac{X_1^v - X_k^v}{\sigma(2t + 2d - 2v)^{\frac{1}{2}}}\right)$$

where

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2} y^2)\, dy.$$

Let the following be our rule: the k^{th} path is cut off at time v if the probability that $X_1^{d+t} \geq X_k^{d+t}$ exceeds $1 - \alpha$. Define ρ such that $\Phi(\rho) = 1 - \alpha$. Thus if

(2) $$X_1^v - X_k^v \geq \rho\cdot(2t + 2d - 2v)^{\frac{1}{2}} \cdot \sigma,$$

then the k^{th} path is stopped at the v^{th} stage, instead of going d stages deep. This approach also corresponds to testing the hypothesis that $X_k^{d+t} \geq X_1^{d+t}$ with a level of significance α (the level of significance is the maximum probability of rejecting the hypothesis when the hypothesis is true). In this case we reject the hypothesis if $X_1^v - X_k^v$ satisfied Eqn. (2).

As the random walk model of Eqn. (1) allows us to cut some paths off in the forward evaluation, there should be a significant reduction in calculations. Note that the estimates for μ and α also provide an indication of the accuracy of the polynomial surface. If μ is positive, then this means that the polynomial estimates tend to be lower than the true values; and if σ^2 is small, then this implies that there is not much variability in the polynomial estimates from one stage to another. When comparing alternative surfaces, the surface with the smallest value of σ^2 is the best according to this criterion. Reference 1 gives extensive empirical results from applying this error analysis approach to a game problem.

5. APPLICATIONS

Next, we will discuss some applications of DYGAM that have been made. The two-piston game is played over the network illustrated in Figure 3, which consists of two combat nodes plus two reserve nodes. The combat nodes can move vertically along parallel and independent tracks. Each side has its own reserve node where new elements are generated during the course of the game. Red elements are not allowed to be at Blue's reserve node, and vice versa. However, both sides may have elements at the combat nodes. Thus, each side can occupy at most three of the four nodes in the network. When both sides have elements at the same combat node, attrition occurs at this node. The positions of the nodes are not fixed, but are allowed to move during the game. For example, if Blue has a stronger force than Red has at a combat node, then Red will be forced to retreat, which corresponds to having the node move backwards into Red territory according to a specified movement formula. The transit time between each reserve and combat node is assumed to remain constant, so that as the combat nodes move, the reserve nodes will adjust accordingly. The game is played over T time periods in a sequential manner; first one side moves, then the other side moves, etc. Each time period could refer to a day, half-day, etc. The payoff to each side at the end of the game is a function of the final positions of the combat nodes and of the number and locations of the remaining elements. The problem faced by each side is how to deploy and manipulate its forces in order to maximize its payoff at the end of the game. The payoff to Red is assumed to be the negative of the payoff to Blue, so that we have a zero-sum game.

The two-piston game consists of eight state variables: the number of Red elements at the first and second combat node and at the reserve; the number of Blue elements at the first and second combat node and at the reserve; and the positions of the two combat nodes. However, Reference 9 showed that under

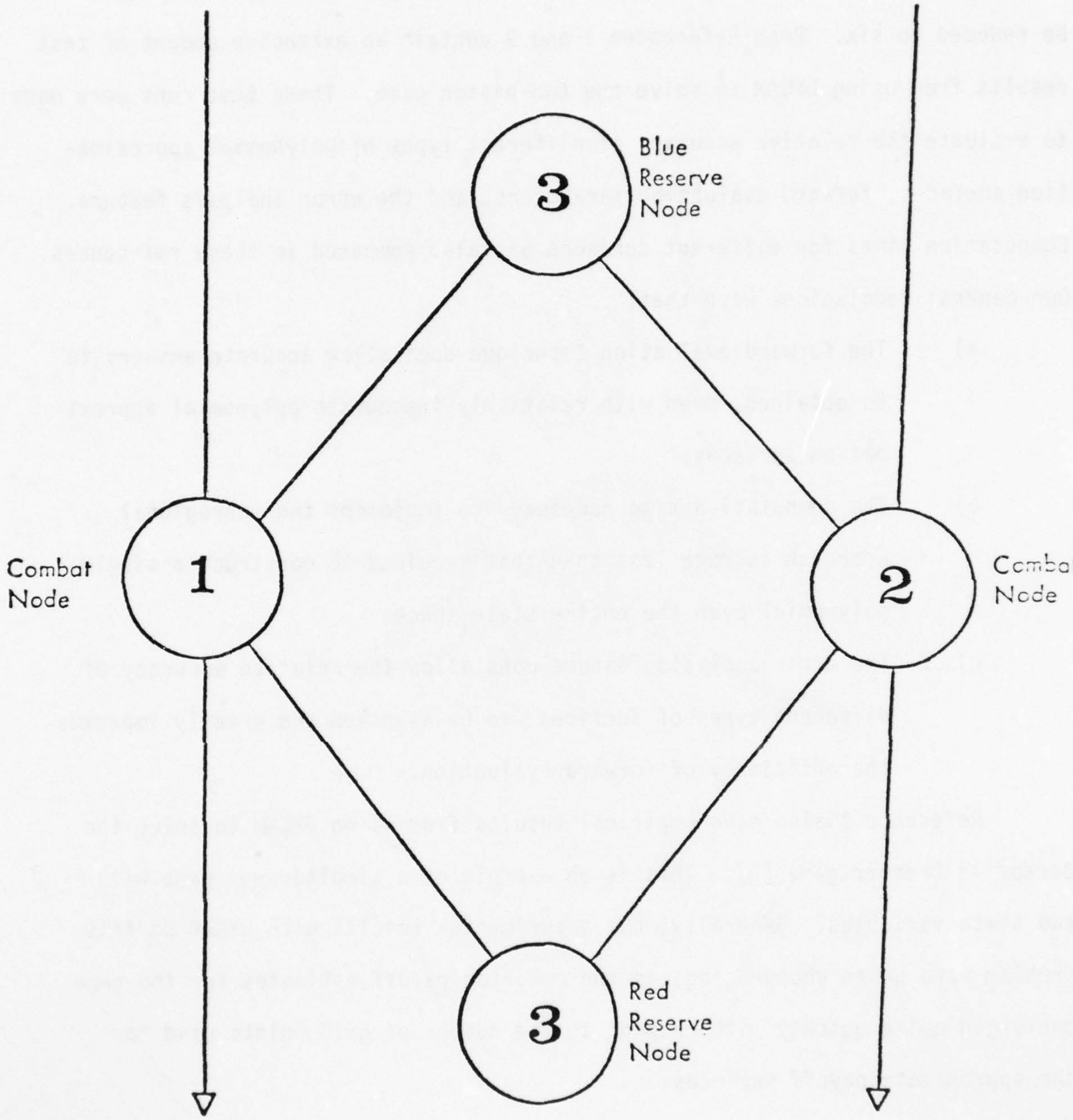


Figure 3. The Two-Piston Game

certain conditions concerning the form of the final payoff function, the attrition, and movement equations, the effective number of state variables could be reduced to six. Both References 1 and 9 contain an extensive amount of test results from using DYGAM to solve the two-piston game. These test runs were made to evaluate the relative accuracy of different types of polynomial approximation surfaces, forward evaluation parameters, and the error analysis feature. Computation times for different surfaces are also compared in these references. Our general conclusions were that:

a) The forward evaluation technique does allow accurate answers to be obtained, even with relatively inaccurate polynomial approximation surfaces;

b) The computation time necessary to implement the subregional approach is much less than that required to construct a single polynomial over the entire state space;

c) The error analysis feature does allow the relative accuracy of different types of surfices to be assessed and greatly improves the efficiency of forward evaluation.

Reference 1 also gave empirical results from using DYGAM to solve the Berkovitz-Dresher game [6]. This is an example of a simultaneous game with two state variables. Generally, out experimental results with DYGAM on this problem were quite encouraging, as the computed payoff estimates for the game converged quite quickly with respect to the number of grid points used for the approximate payoff surfaces.

REFERENCES

1. Z.F. Lansdowne, G.B. Dantzig, R.P. Harvey, and R.D. McKnight, Development of an Algorithm to Solve Multi-Stage Games (prepared for the Office of Studies and Analysis, U.S. Air Force), Control Analysis Corporation, May 1973.

2. J. Bracken, Two Optimal Sortie Allocation Models, Paper P-992, Institute for Defense Analysis, December 1973.

3. J.R. Fish, ATACM: ACDA Tactical Air Campaign Model (prepared for U.S. Arms Control and Disarmament Agency), KETRON, Inc., October 1975.

4. M. Dresher, "The N-Stage Game and Lagrangian Multipliers; the N-Stage Game and DYGAM," Technical Note NWRC-TN-61, Stanford Research Institute, November 1975.

5. L.B. Anderson, et al., On Maxmin and Minmax Strategies in Multi-Stage Games and ATACM, Paper P-1197, Institute for Defense Analysis, August 1976.

6. L. D. Berkovitz and M. Dresher, "A Game Theory Analysis of Tactical Air War," Operations Research, Vol. 7, 1959, pp. 599-820.

7. R.E. Bellman, Dynamic Programming, Princeton University Press, Princeton, 1957.

8. R.E. Bellman and S.E. Dreyfus, Applied Dynamic Programming, Princeton University Press, Princeton, 1962.

9. Approximate Solutions of Multi-Stage Network Games, Control Analysis Corporation, August 1972.

APPENDIX B

DESCRIPTION OF THE SOLUTION PROCEDURE USED IN ATACM*

by
F. A. Miercort

Presented to the Theater-Level
Gaming and Analysis Workshop
For Military Force Planning
(27-29 September 1977)

*The material in this paper is excerpted from "ATACM: ACDA Tactical Air Campaign Model," by John R. Fish, Ketron, Inc., October 1975.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

The ACDA Tactical Air Campaign Model (ATACM) is a computer model designed and built for the Arms Control and Disarmament Agency (ACDA) for use in analyzing the impacts of various Mutual Balanced Force Reduction (MBFR) proposals upon a tactical airwar in Europe between NATO and Warsaw Pact forces. The design of ATACM was based upon the findings of a survey (Reference 1) of existing tactical air models conducted by KETRON to assess the applicability of existing models to ACDA's requirements. Results of the survey indicated the need for a new model incorporating the most desirable features of existing models (e.g., TAC CONTENDER, VECTOR, OPTSA, DYGAM) into a rigorous optimization framework allowing more aircraft types and a wider selection of aircraft missions.

As a realization of the survey's recommendations, ATACM models an air campaign as a zero-sum staged game between opposing forces and employs dynamic programming to solve this game for approximate, optimal aircraft allocation strategies for both sides at each stage of the campaign. Because of the economies associated with the optimization methodology used, ATACM offers many features previously not practical in other optimizing models. Specifically, ATACM permits:

- as many as four user-defined aircraft types per side and as many as eight different missions per aircraft type
- automatic generation of approximate, optimal, enforceable aircraft allocation strategies as a function of stage for any subset of the missions for which user-specified fractions are not supplied. The user may specify all, part, or none of the allocation fractions for a given aircraft type and the model generates optimal values for those fractions not specified.
- calculation of firm upper and lower bounds on the objective function value associated with the enforceable strategies employed
- option to use a weighted sum of three different objective functions as the criterion for generating the optimal strategies
- option to individually weight the Blue and Red contributions to these objective functions as a function of stage
- option to specify fractional or numerical reinforcements for any aircraft type as a function of stage

Following sections present a description of how the problem of tactical air warfare is formulated in ATACM, a description of the attrition relationships used to evaluate the outcomes of air-to-air, air-to-ground, and ground-to-air engagements, and an outline of the optimization methodology used to generate optimal enforceable strategies and objective function bounds.

PROBLEM FORMULATION

ATACM formulates a tactical air campaign as a staged game between opposing Blue and Red air and ground forces. It generates for each side, and for each stage of the war, strategies which optimize the utilization of these forces over the length of the campaign. Figure 1 presents a graphical representation of a general staged game which depicts the roles of those essential elements which will be described in detail in following sections.

In Figure 1, for each stage or time period of the campaign, vertical planes represent the state space of possible beginning and ending force levels for the opposing sides. Corresponding to each point in the beginning state space for a given stage, a game matrix can be constructed with m_t and n_t strategies available to Blue and Red respectively. The numbers of strategies, m_t and n_t, are a function of stage, while the one-stage payoffs in the game matrix depend upon the starting resource levels, the objective function chosen as a measure of overall performance, and the strategies selected by both sides.

For a given strategy selection at the beginning of stage 1, assessment relationships determine the value of the payoff and the attrition or losses suffered by both sides as a result of the one-stage battle. The dashed line in Figure 1 from stage 1 to stage 2 depicts the effect of this attrition and shows that the starting force levels for stage 2 will, barring reinforcements, generally be less than those for stage 1. The decision facing both sides at stage 2 is analogous to that at stage 1, the only possible difference being the strategies available to each side and the associated one-stage payoffs. Once again strategies are chosen, the objective function is incremented by the corresponding payoff, attrition relationships map the starting force level for the current stage into a new point in the state space for the next stage, etc. This process is repeated for the number of stages in the game, and the value of the objective function summed over all stages determines the outcome of the campaign.

Within the general framework depicted in Figure 1, the key elements in the formulation of a tactical air war as a staged game are the forces, strategies, and objective functions. The way these elements are defined in ATACM is described in the following paragraphs of this section. The key elements in the solution of a tactical air war include the assessment methodology, used to compute payoffs and attrition, and the optimization methodology used to select optimal strategies at each stage of the campaign. These elements are discussed in following sections.

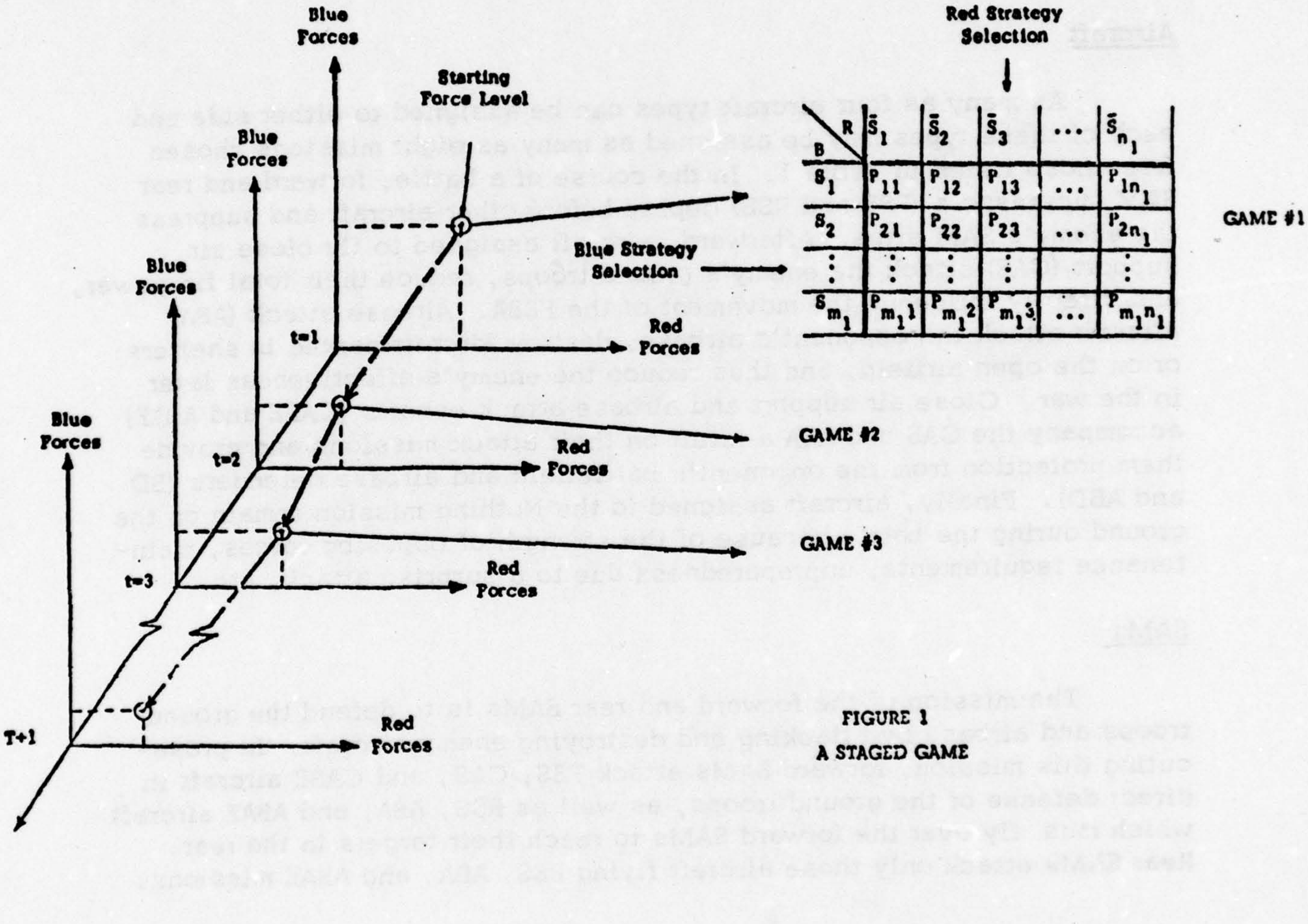


FIGURE 1
A STAGED GAME

OPPOSING FORCES AND MISSIONS

In the ATACM formulation, opposing forces consist of aircraft, SAMs, and ground divisions deployed in the stylized scenario depicted in Figure 2. A single airbase on each side serves as the staging area for all air missions flown against the opponent's SAMs, ground troops, or airbase. SAMs, which may be interpreted as any type of surface-to-air defense weapons, are segregated into forward and rear components corresponding to the location of the area they defend. Ground troops are defined in terms of homogenous divisions fighting on either side of a single-sector front (FEBA).

Aircraft

As many as four aircraft types can be assigned to either side and each of these types may be assigned as many as eight missions chosen from those listed in Table 1. In the course of a battle, forward and rear SAM suppressors (FSS and RSS) deploy before other aircraft and suppress the enemy's SAM sites. Afterward, aircraft assigned to fly close air support (CAS) attack the enemy's ground troops, reduce their total firepower, and directly influence the movement of the FEBA. Airbase attack (ABA) aircraft attack the opponent's airbase, destroy aircraft parked in shelters or on the open airfield, and thus reduce the enemy's effectiveness later in the war. Close air support and airbase attack escorts (CASE and ABAE) accompany the CAS and ABA aircraft on their attack missions and provide them protection from the opponent's battlefield and airbase defenders (BD and ABD). Finally, aircraft assigned to the Nothing mission remain on the ground during the battle because of the strength of opposing forces, maintenance requirements, unpreparedness due to a surprise attack, etc.

SAMs

The mission of the forward and rear SAMs is to defend the ground troops and airbase by attacking and destroying enemy aircraft. In prosecuting this mission, forward SAMs attack FSS, CAS, and CASE aircraft in direct defense of the ground troops, as well as RSS, ABA, and ABAE aircraft which must fly over the forward SAMs to reach their targets in the rear. Rear SAMs attack only those aircraft flying RSS, ABA, and ABAE missions.

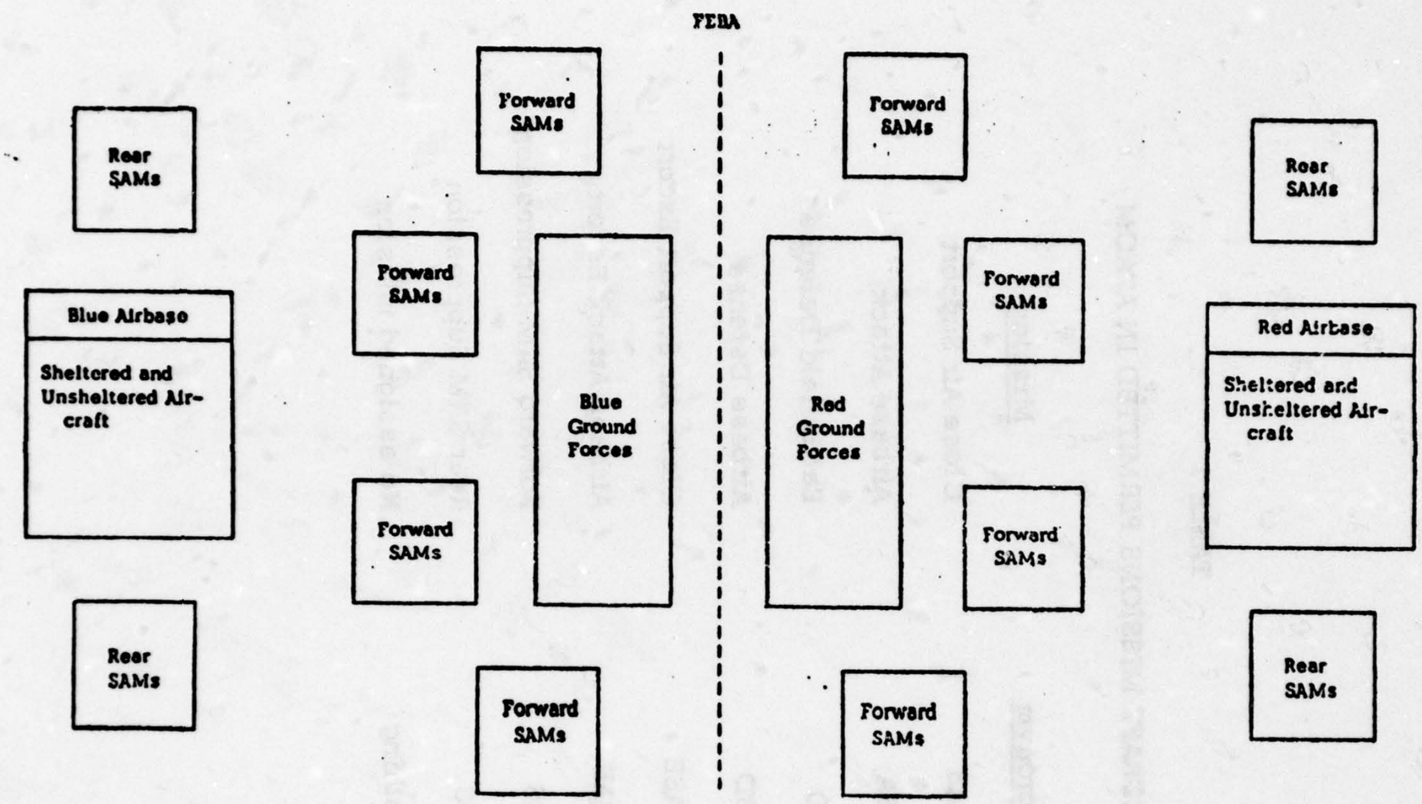


FIGURE 2

ATACM SCENARIO

TABLE 1

AIRCRAFT MISSIONS PERMITTED IN ATACM

Acronym	Mission
CAS	Close Air Support
ABA	Airbase Attack
BD	Battlefield Defense
ABD	Airbase Defense
CASE	Close Air Support Escort
ABAE	Airbase Attack Escort
FSS	Forward SAM Suppression
RSS	Rear SAM Suppression
Nothing	No assigned mission

Ground Troops

Because ATACM was designed as a tool for studying the effects of different numbers and types of aircraft upon the outcome of an air campaign, the ground representation is simplistic and serves primarily as an input to the figure of merit used in the optimization process. Ground troops are segregated into homogeneous divisions with each division assigned a maximum firepower score. Successful CAS sorties flown against the enemy troops reduce this maximum firepower by a specified amount, and FEBA movement is then calculated as a user-specified function of the ratio of the total net firepowers delivered by Blue and Red respectively.

STRATEGIES

Given the forces and missions described above, the strategy a commander uses during a given time period or stage is a fractional allocation of all aircraft to missions. For example, if only one aircraft type is available to the Blue commander, and this aircraft can prosecute four missions selected from Table 1, the set of possible strategies from which he may choose can be characterized as the set of all 4-tuples whose elements are positive fractions which sum to one. Examples include (.5, 0, .5, 0), (.5, 0, .2, .3), (.25, .25, .25, .25), etc. In the general case of s missions, s-tuples representing possible fractional allocations for one aircraft type are called <u>decision vectors</u>.

If the Blue commander has only one aircraft type, the sets of possible decision vectors and strategies are identical. If two aircraft types are available to the Blue side, the set of possible strategies corresponds to the set of all possible decision vector pairs with the first decision vector representing allocations for aircraft type 1, the second allocations for aircraft type 2. An example of a strategy for two aircraft types, each with four possible missions, would be

$$\big((.5,\ 0,\ .5,\ 0),\ (.5,\ .2,\ .2,\ .1)\big)$$

Analogously, possible strategies for a side with three aircraft types can be represented as decision vector triples, etc.

To limit the number of decision vectors (and thus strategies) from which ATACM must choose an optimal allocation, a parameter called a <u>minimum allocation fraction</u> is specified for each aircraft type. The minimum allocation fraction for an aircraft type is the smallest fractional unit which can be assigned to any mission. In the case of an aircraft type with four

assigned missions, a minimum allocation fraction of .5 limits the set of possible decision vectors to the following ten.

(1, 0, 0, 0)	(0, .5, 0, .5)
(.5, 0, 0, .5)	(0, .5, .5, 0)
(.5, 0, .5, 0)	(0, 0, 1, 0)
(.5, .5, 0, 0)	(0, 0, .5, .5)
(0, 1, 0, 0)	(0, 0, 0, 1)

Correspondingly, the number of strategies available to a side with two, three, or four such aircraft types would be 100, 1000, or 10,000 respectively. In general, the number of possible decision vectors V for an aircraft type with s missions and a minimum allocation fraction equal to 1/t is given by

$$V = \frac{(s + t - 1)!}{(s - 1)!\, t!} \tag{1}$$

In addition to aircraft types, missions assigned, and minimum allocation fractions, ATACM permits one other important specification which determines the set of strategies available for a particular stage of the conflict. The user is allowed to specify, for any stage, a fixed assignment of aircraft to mission in terms of a multiple of the minimum allocation fraction. Looking at the previous example, the user can force half of the aircraft to prosecute the first mission assigned by specifying the first element in the corresponding decision vector to always equal .5. In that case, the set of possible decision vectors for the specified stage and aircraft type reduces to the following subset of those shown above:

(.5, 0, 0, .5)
(.5, 0, .5, 0)
(.5, .5, 0, 0)

The set of all strategies generated from these three decision vectors will reflect the specified allocation, and the strategy selected from this set by ATACM will optimize remaining allocations over those missions for which fractions are not specified.

As can be seen from this example, as more and more fractions are specified, the number of possible decision vectors and strategies decreases, and the process of strategy selection optimizes over fewer and fewer missions. In the extreme case, where all fractions in all decision vectors are specified, the set of strategies available at each stage reduces to a single strategy. Strategy selection then becomes a vacuous operation, and the net effect is an evaluation of an air campaign using a user-specified strategy at each

stage. Thus, depending upon the number of mission allocations specified, ATACM can be used as an optimization, sub-optimization, or strategy-specified model.

OBJECTIVE FUNCTIONS

ATACM permits the user to select any linear combination of three objective functions to be used as the overall measure of the opposing forces' performance during an air campaign. Specifically, the overall objective function used as the criterion for strategy selection can be expressed as

$$F = w_1 f_1 + w_2 f_2 + w_3 f_3 \tag{2}$$

where f_1 = difference of total Blue minus total Red CAS firepower

f_2 = difference of total Blue minus total Red (CAS firepower + ground firepower)

f_3 = total FEBA movement computed as a user-specified function of the ratio of Blue's total ground firepower to Red's

and w_j = user specified weight on f_j, $j = 1, 2$, or 3.

By appropriate choice of the w_j, the user can optimize using any one of the f_j, or he can generate hedging strategies -- those not precisely optimal for any single criterion but instead optimal for several criteria at once -- by specifying F to be a combination of the f_j. Regardless of what F is specified, f_1, f_2, and f_3 are also computed and recorded individuaaly making it possible to simultaneously monitor the effects of different optimization criteria on each of the objective functions.

In addition to the weights w_1, w_2, and w_3, ATACM also permits the user to weight the Blue and Red components of f_1, f_2, and f_3 by stage. To illustrate, each f_j can be expanded as follows:

$$f_1 = \sum_{t=1}^{T+1} b_t CAS_{Bt} - r_t CAS_{Rt} \tag{3}$$

$$f_2 = \sum_{t=1}^{T+1} b_t TFP_{Bt} - r_t TFP_{Rt} \tag{4}$$

$$f_3 = \sum_{t=1}^{T+1} \frac{b_t + r_t}{2} \; FEBA_t \qquad (5)$$

where T = number of stages in the campaign

$$CAS_{kt} = \begin{cases} \text{CAS firepower delivered by side k during stage t} & \text{for } t \leq T \\ \text{residual value of undamaged aircraft on side k at end of war} & \text{for } t+T+1 \end{cases}$$

$$TFP_{kt} = \begin{cases} \text{total firepower (ground + CAS) delivered by side k during stage t} & \text{for } t \leq T \\ \text{residual value of undamaged aircraft on side k at end of war} & \text{for } t+T+1 \end{cases}$$

$$FEBA_t = \begin{cases} \text{FEBA movement during stage t} & \text{for } t \leq T \\ 0 & \text{for } t=T+1 \end{cases}$$

b_t = weight on Blue's contribution to the objective function during stage t ($b_{T+1}=1$)

r_t = weight on Red's contribution to the objective function during stage t ($r_{T+1}=1$)

Depending upon the scenario being simulated, the weights b_t and r_t can be used to reflect the various effects of logistics, force readiness, pilot skills, ground terrain, etc., all as a function of stage. For example, in the case of a surprise attack by Red, the amount of firepower Blue can deliver per sortie during early stages of the war might be severely limited by logistics, readiness, etc. To simulate this situation, weights b_t specified for the first few stages of the war would be smaller than those specified for later stages.

ASSESSMENT METHODOLOGY

Returning to the general description of a staged game depicted in Figure 1, the assessment methodology computes the values of the payoffs in each game matrix and the attrition suffered by both sides as a result of each possible one-stage battle. The important considerations in describing this methodology include the types of engagements which may occur between the opposing sides, the sequence in which these engagements occur within each stage, and the relationships used to compute attrition for each type of engagement. An important finding of the survey of existing models was the general lack of agreement concerning the best way to treat each of these facets of the assessment procedure. The assessment methodology described below is a mix of those used in OPTSA and VECTOR (References 2 and 3) and consequently suffers from some of the same limitations cited in Reference 1. In consideration of these limitations, ATACM is purposely structured so that other, alternative assessment methodologies can be implemented with minimal programming effort.

TYPES OF ENGAGEMENTS

In the current version of ATACM the types of engagements permitted can be classified as air-to-air, air-to-ground, and ground-to-air. Air-to-air engagements occur when CAS and CASE aircraft engage battlefield defenders or when ABA and ABAE aircraft engage airbase defenders. Air-to-ground engagements occur when ABA aircraft attack the opponent's airbase, when SAM suppressors attack the opponent's SAMs, or when CAS aircraft deliver ordnance against the opponent's ground troops. Ground-to-air engagements occur between SAMs and opposing aircraft flying SAM suppression, CAS, CASE, ABA, or ABAE missions.

ENGAGEMENT CYCLES

The first event in each stage of the air campaign is the addition of any numerical or fractional aircraft reinforcements specified by the user. Thereafter, attrition and payoffs for each strategy pair are computed and accumulated over a specified number of equal length time periods (e.g. days) called engagement cycles. The first event in each engagement cycle is the assignment of aircraft to missions according to the strategy pair being examined. Then, using specified sortie rates per cycle, these

assignments are converted into sorties which progress en masse through the engagements described below.

FSS and RSS Engagements

FSS and RSS missions depicted in Figure 3 are flown by each side before all other missions in an attempt to clear corridors for subsequent battlefield and airbase attackers. SAMs successfully attacked by suppressors are killed for the duration of the stage in which they are suppressed, but are replaced or restored at the beginning of the next stage. SAM suppressors successfully engaged by SAMs or destroyed by the opponent's airbase attackers are lost for the duration of the campaign.

CAS, CASE, and BD Engagements

The CAS, CASE, and opposing BD missions depicted in Figure 4 are flown after the SAM suppressors. All air-to-air engagements between CAS, CASE, and the opponent's BD sorties are one-on-one encounters. Excess attackers or defenders not engaged in one phase of the engagement cycle proceed unmolested to the next phase. Sorties which are engaged abort their original mission, fight the opposing aircraft, and, if not killed, return to their own airbase. Each successful CAS sortie delivers ordnance on the opponent's ground troops and reduces their total firepower (computed at the beginning of each stage as number of divisions times firepower per division) by a user-specified amount. Thus ground troops, like SAMs, can be thought of as being killed for the duration of the stage in which they are attacked, but replaced at the beginning of the next stage. CAS, CASE, or BD killed in air-to-air engagements or destroyed by the opponent's airbase attackers are lost for the duration of the campaign.

ABA, ABAE, and ABD Engagements

The ABA, ABAE, and opposing ABD missions depicted in Figure 5 are the last missions flown in each engagement cycle. All air-to-air engagements between ABA, ABAE, and the opponent's ABD sorties are one-on-one encounters treated in the same way as those described for CAS, CASE, and BD missions. Each successful ABA sortie delivers ordnance on the opponent's airbase destroying those sheltered and unsheltered aircraft specified as being vulnerable to airbase attack. Shelters are assigned to vulnerable aircraft types in proportion to their relative numbers. Shelters are not destroyed; damaged shelters are

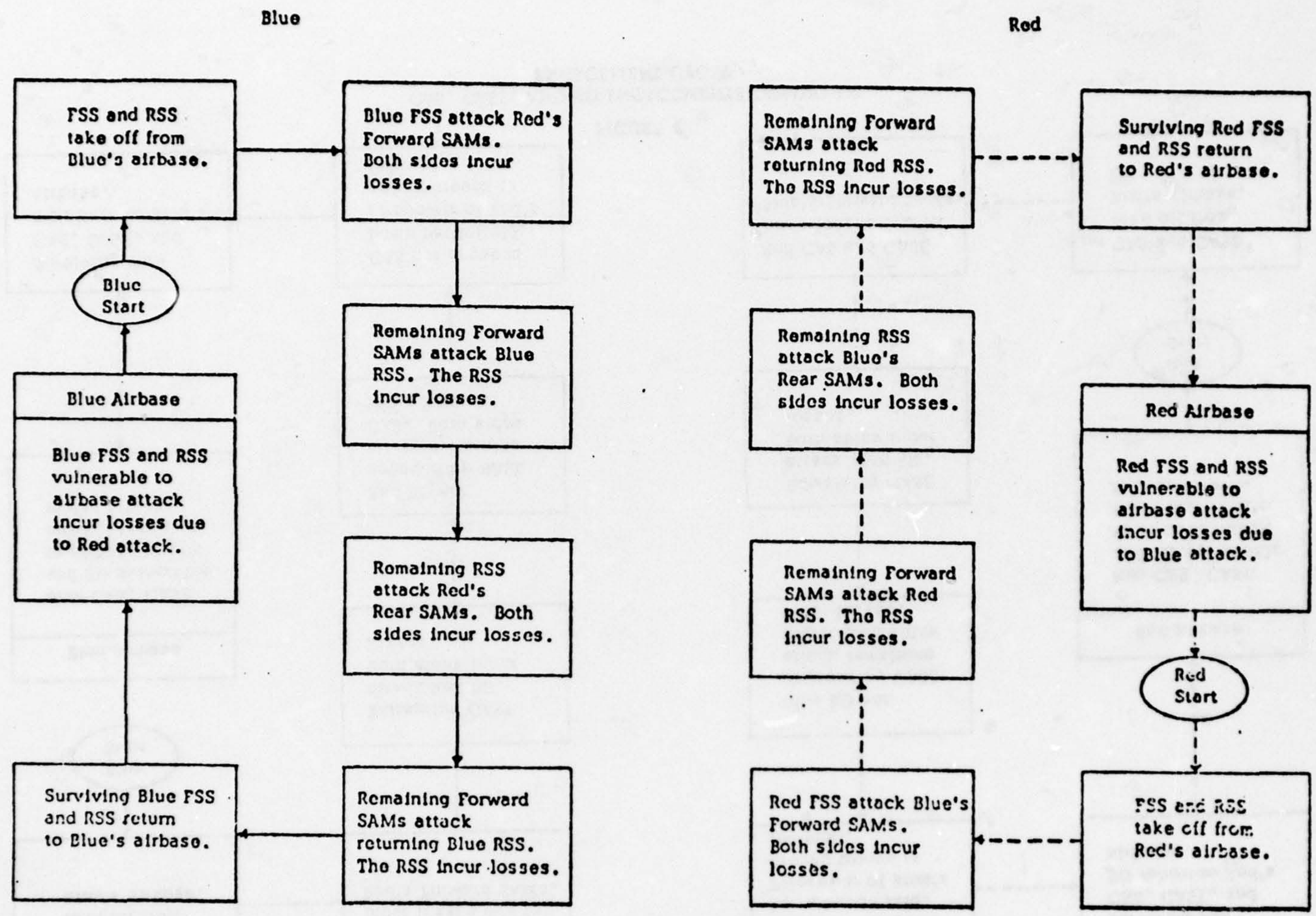


FIGURE 3

FSS AND RSS ENGAGEMENTS DURING AN ENGAGEMENT CYCLE

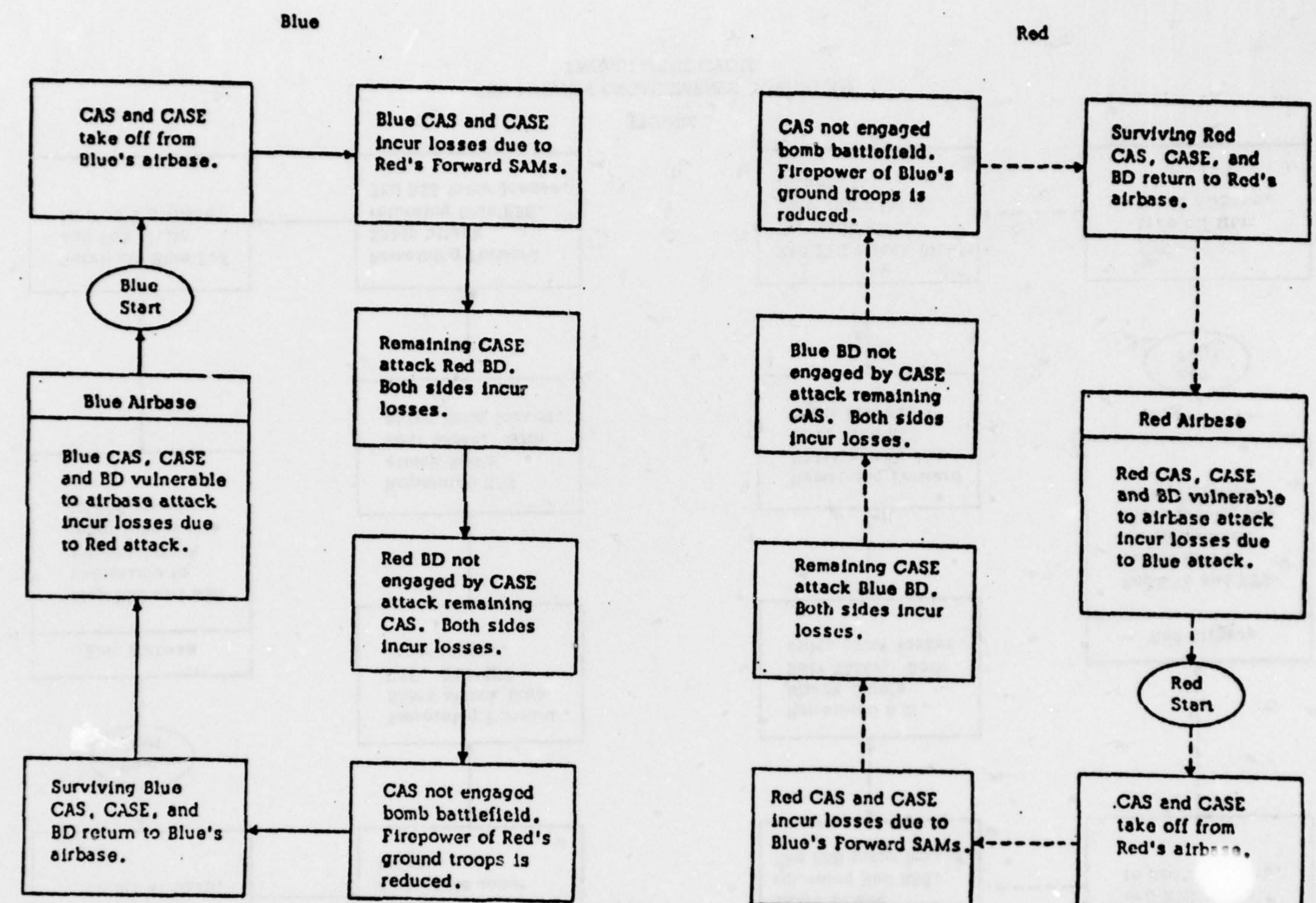


FIGURE 4
CAS, CASE, AND BD ENGAGEMENTS DURING AN ENGAGEMENT CYCLE

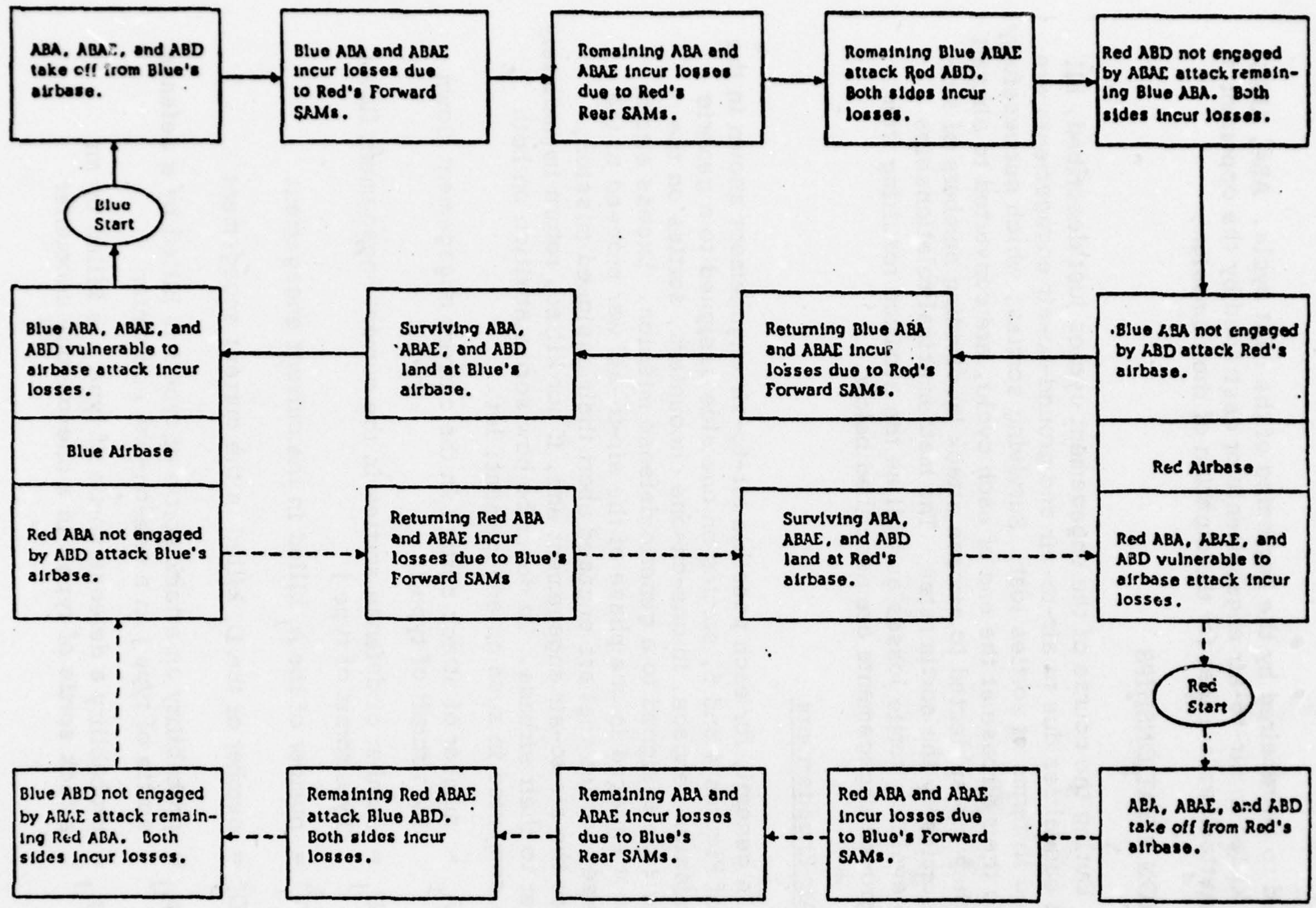


FIGURE 5

ABA, ABAE, AND ABD ENGAGEMENTS DURING AN ENGAGEMENT CYCLE

assumed to be repaired by the beginning of the next cycle. ABA, ABAE, or ABD killed in air-to-air engagements or destroyed by the opponent's airbase attackers are lost for the duration of the campaign.

ATTRITION RELATIONSHIPS

During the course of the engagement cycles just described, all aircraft casualties due to air-to-air and ground-to-air engagements are computed in terms of sorties lost. Surviving sorties, which successfully return to their airbase at the end of each cycle, are converted to aircraft before they are subjected to airbase attack by dividing numbers of surviving sorties by the sortie rates. The mathematical relationships used to evaluate sortie losses as well as the attrition resulting from air-to-ground engagements are described below.

Air-to-Air Engagements

In general, for each possible air-to-air engagement shown in the blocks of Figures 4 and 5, sorties on one side assigned to a generic attack mission, engage, in one-on-one encounters, sorties on the opposing side assigned to a generic defense mission. Excess attack sorties not engaged in one phase of the air-to-air war proceed to the next phase; sorties that are engaged abort their assigned mission, prosecute the air-to-air engagement and, if not killed, return immediately thereafter to their airbase. To describe how sortie attrition on both sides is computed in such an engagement, let

A_i = number of attack sorties in the current engagement flown by aircraft of type i

D_j = number of defense sorties in the current engagement flown by aircraft of type j

A_i^k = number of the A_i killed in the current engagement

D_j^k = number of the D_j killed in the current engagement

p_{ij} = probability an attack sortie of type i is killed by a defense sortie of type j in a one-on-one encounter

q_{ji} = probability a defense sortie of type j is killed by an attack sortie of type i in a one-on-one encounter

To compute A_i^k and D_j^k, the total numbers of attack and defense sorties are first used to compute E, the number of one-on-one encounters, as

$$E = \min \left(\sum_i A_i, \sum_j D_j\right) \tag{6}$$

The allocation of this total to individual aircraft types is then computed proportionally as

$$E_{ij} = E \left(\frac{A_i}{\sum_i A_i}\right)\left(\frac{D_j}{\sum_j D_j}\right) \tag{7}$$

where E_{ij} represents the number of one-on-one encounters between attackers of type i and defenders of type j.

Finally, expected numbers of killed attack and defense sorties flown by aircraft of types i and j respectively are computed as

$$A_i^k = \sum_j E_{ij}\, p_{ij} \tag{8}$$

and

$$D_j^k = \sum_i E_{ij}\, q_{ji} \tag{9}$$

Air-to-Ground Engagements

Of the three types of air-to-ground engagements considered in ATACM, only the SAM suppression and airbase attack missions directly produce losses among the opponent's forces. As described in the discussion of Figure 4, CAS sorties reduce the opponent's ground firepower by a simple subtractive rule which indirectly reflects ground losses. By contrast, SAM and parked aircraft losses are computed using exponential relationships relating the kill probabilities and numbers of attackers (SAM suppressors or ABA) to the number of opponents (SAMs or parked aircraft).

SAM Losses

To derive the expression for SAM losses produced by generic (FSS or RSS) SAM suppression sorties, let

A_i = number of SAM suppression sorties in the current engagement flown by aircraft of type i

D = number of opposing SAMs

D^k = number of D killed in the current engagement

q_i = probability a SAM is killed by a SAM suppressor sortie of type i

To compute D^k, the numbers of SAM suppression sorties flown are used as weights to compute an average probability of kill

$$\bar{q} = \frac{\sum_i A_i q_i}{\sum_i A_i} \tag{10}$$

This probability, along with the numbers of suppression sorties and SAMs, is used to compute D^k as

$$D^k = D\left(1 - \exp\left(-\bar{q}\,\frac{\sum_i A_i}{D}\right)\right) \tag{11}$$

Losses Due to ABA

The attrition relationship for computing the effect of ABA sorties on parked aircraft is analogous to Equation (11), the only difference being the number of opponent types. Specifically, let -

A_i = number of ABA sorties in the current engagement flown by aircraft of type i

D_j = number of parked aircraft of type j vulnerable to airbase attack. Vulnerable aircraft are assigned to shelters in proportion to their relative numbers in the inventory; j=1 corresponds to sheltered aircraft, j=2 to unsheltered.

D_j^k = number of D_j killed in the current engagement.

q_{ji} = probability a vulnerable aircraft of type j is killed by a sortie flown by an ABA aircraft of type i

To determine D_j^k, first E_{ij}, the number of attack sorties of each type i assumed to attack vulnerable aircraft of type j, is computed proportionally as

$$E_{ij} = A_i \frac{D_j}{\sum_j D_j} \tag{12}$$

Using the E_{ij} as weights, an average probability of killing a parked aircraft of type j is computed as

$$\bar{q}_j = \frac{\sum_i E_{ij} q_{ji}}{\sum_i E_{ij}} \tag{13}$$

Finally, D_j^k is computed using the standard exponential expression

$$D_j^k = D_j \left(1 - \exp\left(- \bar{q}_j \frac{\sum_i E_{ij}}{D_j} \right) \right) \tag{14}$$

Ground-to-Air Engagements

In ground-to-air engagements, SAMs attack opposing sorties in one-on-one encounters analogous to one-sided air-to-air engagements. If the number of opposing sorties exceeds the number of SAMs, excessive sorties are not attacked. If the number of SAMs exceeds the number of sorties, excessive SAMs are not launched. To describe the attrition produced by a generic SAM (forward or rear) attack, let

A = number of SAMs in the current engagement

D_j = number of target sorties in the current engagement flown by aircraft of type j

D_j^k = number of D_j killed in the current engagement

q_j = probability a target sortie flown by an aircraft of type j is killed by a SAM in a one-on-one encounter

To compute D_j^k, the total number of SAMs and opposing aircraft sorties are used to compute E, the number of one-on-one encounters, as

$$E = \min(A, \sum_j D_j) \tag{15}$$

$$E_j = E \frac{D_j}{\sum_j D_j} \tag{16}$$

where E_j represents the number of one-on-one encounters between SAMs and target sorties of type j.

Finally, the expected number of killed sorties flown by aircraft of type j is computed as

$$D_j^k = E_j q_j \tag{17}$$

OPTIMIZATION METHODOLOGY

The methodology used to select optimal strategies for both sides at each stage of the campaign reflects the recommendations of the survey presented in Reference 1. The basic approach is to select enforceable strategies which individually optimize each side's minimal performance over the length of the campaign. Following paragraphs explain this optimization criterion in detail and describe the dynamic programming methodology used to implement it.

MAXMIN/MINMAX STRATEGIES AND PAYOFFS

In the standard game matrix such as that shown in Figure 1 the possible objective functions used to compute payoffs to Blue are defined such that positive payoffs indicate Blue success. Blue's objective is to choose that strategy which will produce the largest payoff, while Red's objective is to choose that strategy which will produce the smallest payoff. The approach used in ATACM is to assume both Blue and Red act conservatively in choosing their strategies. To illustrate, for a simple case, Figure 6 presents a game matrix for starting force levels in a hypothetical one-stage campaign in which Blue has five possible strategies, Red has four.

One-Stage Game

Since Blue and Red are assumed to be conservative, each chooses that strategy which will produce the most favorable outcome under the worst of circumstances -- i.e. prior knowledge of his selection by his opponent. In the case of Blue, selection of S_1 guarantees a payoff no less than 1 regardless of which strategy Red chooses. Selection of S_2 guarantees a payoff no less than 0, S_3 no less than -1, etc. These minimal payoffs (or row minimums) are shown in Figure 6 for each possible Blue selection. Assuming Red has superior intelligence, Blue would choose that strategy, S_1, which maximizes over the set of minimal payoffs. Thus S_1 is called Blue's MAXMIN strategy, and 1, the minimal payoff associated with S_1, is called the MAXMIN payoff or objective function value. The Red strategy $\bar{S}_2$ which would have to be play against S_1 to yield the MAXMIN payoff is called Red's MAXMIN strategy.

	Blue \ Red	$\bar{S}_1$	$\bar{S}_2$ (Red MAXMIN)	$\bar{S}_3$	$\bar{S}_4$ (Red MINMAX)	Row Min
Blue MAXMIN →	S_1	2	$\boxed{1}$	4	2	1
	S_2	3	5	0	0	0
Blue MINMAX →	S_3	6	-1	4	$\boxed{3}$	-1
	S_4	3	1	2	-2	-2
	S_5	0	-3	-5	1	-5
	Col Max	6	5	4	3	

FIGURE 6

ONE-STAGE GAME MATRIX

In the case of Red's selection, choice of $\bar{S}_1$ could result in a payoff as large as 6 if Blue were to play S_3, selection of $\bar{S}_2$ could result in a payoff as large as 5, $\bar{S}_3$ a payoff as large 4, and $\bar{S}_4$ a payoff as large as 3. Assuming Blue has superior intelligence, conservative Red would choose that strategy, $\bar{S}_4$, which minimizes over the set of column maximums shown in Figure 6. Strategy $\bar{S}_4$ is called Red's MINMAX strategy, S_3 is called Blue's MINMAX strategy, and 3, the payoff associated with playing $\bar{S}_4$ against S_3 is called the MINMAX payoff.

If Blue were to play its MAXMIN strategy and Red its MINMAX strategy, the payoff which would result is 2. This MAXMIN vs. MINMAX payoff will always be greater than or equal to the MAXMIN payoff (in this case 1) and less than or equal to the MINMAX payoff (in this case 3).

From this example, it should be clear how the optimization criterion used in ATACM would be applied for a one-stage campaign. Given the starting force levels, strategies, and objective function specification, payoffs in the corresponding game matrix would be computed using the assessment methodology described in the previous section. Blue's MAXMIN strategy and payoff would be computed by maximizing over row minimums while Red's MINMAX strategy and payoff would be computed by minimizing over column maximums. Output would consist of the Blue MAXMIN strategy, the Red MINMAX strategy, the lower and upper MAXMIN/MINMAX objective function bounds, and the actual MAXMIN vs. MINMAX payoff which would result if both sides played their conservative strategies.

Multi-Stage Game

To understand how the MAXMIN/MINMAX strategy selection procedures for a one-stage game extend to a multi-stage game, it is helpful to present an alternative representation of a staged game called an extended game as shown in Figure 7. In an extended game representation of a T-stage air campaign, Blue extended strategies are T-tuples whose t^{th} element is a strategy selected from the set of one-stage strategies $S_1, S_2, \ldots, S_{m_t}$ available to Blue at the t^{th} stage of the campaign. In other words, the first element in an extended strategy for Blue is the strategy Blue would use for stage 1, the 2nd element is Blue's strategy for stage 2, ..., and the T^{th} element is Blue's strategy for the last stage of the campaign. The number of possible extended

FIGURE 7

EXTENDED GAME REPRESENTATION OF A T-STAGE GAME

Red / Blue	$\overline{ES}_1$	$\overline{ES}_2$	...	$\overline{ES}_N$
ES_1	TP_{11}	TP_{12}	...	TP_{1N}
ES_2	TP_{21}	TP_{22}	...	TP_{2N}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
ES_M	TP_{M1}	TP_{M2}	...	TP_{MN}

ES_i = Blue Extended Strategy = $(S_{i_1}, S_{i_2}, \ldots, S_{i_T})$

where S_{i_t} = Blue strategy for stage t

$\overline{ES}_j$ = Red Extended Strategy = $(\bar{S}_{j_1}, \bar{S}_{j_2}, \ldots, \bar{S}_{j_T})$

where $\bar{S}_{j_t}$ = Red strategy for stage t

$M = m_1 \cdot m_2 \cdot \ldots \cdot m_T$ $\qquad$ $N = n_1 \cdot n_2 \cdot \ldots \cdot n_T$

TP_{ij} = Total Payoff produced by playing ES_i against $\overline{ES}_j$

strategies, M, from which Blue may choose is equal to

$$M = m_1 \cdot m_2 \cdot m_3 \cdot \ldots \cdot m_T \tag{18}$$

where m_t equals the number of possible Blue strategies available for stage t. Extended strategies for Red are analogous.

Payoffs in an extended game are <u>total payoffs</u> over the length of the campaign which result from playing a Blue extended strategy against a Red extended strategy. Looking back at Figure 1, a total payoff in the extended game depends upon the one-stage payoffs and the attrition represented by the dashed line. Each different combination of Blue-Red extended strategies corresponds to a unique attrition path from the state space at the beginning of stage 1 to the state space at the end of stage T.

Theoretically, selection of MAXMIN/MINMAX strategies in a multi-stage campaign, represented in the context of the extended game, is exactly the same as that described for a one-stage game. Indeed, for a one-stage campaign the extended game representation reduces to a one-stage game. Unfortunately, for air campaigns with reasonable numbers of strategies and stages the extended game representation is valuable only as an abstraction because of its prohibitive size. For example, if Blue and Red each had only 20 possible strategies to choose from at each stage of a three stage campaign, the corresponding extended game would be an 8000 by 8000 matrix requiring 64 million three-stage payoff evaluations. Assuming 10^{-3} seconds of computer time was required for each one-stage assessment, evaluation of the payoffs alone would require over 50 hours. Needless to say, a method other than the straightforward solution of the extended game is required to determine MAXMIN/MINMAX strategies for the general multi-stage campaign.

DYNAMIC PROGRAMMING SOLUTION

The reason the extended game representation of modest sized air campaigns becomes untractable is that it treats the problem of strategy selection for all stages as a single step process. An alternative approach, using dynamic programming, is to decompose the problem into a series of one-stage problems, similar in many respects to the way the problem was originally posed in Figure 1. To illustrate the dynamic programming approach, its necessary to define more precisely many of the

concepts first presented there. For purposes of illustration, we will assume Blue and Red each have one aircraft type -- more types simply increase the dimensionality of the state space. For this case, shown in Figure 8, let

X_t = a point, (B_t, R_t), in the state space at the beginning of stage t corresponding to B_t and R_t aircraft available to Blue and Red respectively

$S(X_t)$ = Blue's one-stage MAXMIN strategy selection corresponding to X_t

$\bar{S}(X_t)$ = Red's one-stage MINMAX strategy selection corresponding to X_t

$TP(X_t)$ = the total MAXMIN payoff associated with optimal play by Blue from state X_t at the beginning of stage t to the end of the campaign

$\overline{TP}(X_t)$ = the total MINMAX payoff associated with optimal play by Red from state X_t at the beginning of stage t to the end of the campaign

$P_{ij}(X_t)$ = payoff in the one-stage game matrix corresponding to selection of the i^{th} and j^{th} strategies when in state X_t

$Z_{ij}(X_t)$ = state X_{t+1} into which selection of the i^{th} and j^{th} strategies by Blue and Red maps the state X_t

Using these definitions, $TP(X_1)$ and $\overline{TP}(X_1)$ are the desired MAXMIN and MINMAX payoffs associated with the extended strategies

$$ES(X_1) = (S(X_1), S(X_2), S(X_3), \ldots, S(X_T)) \tag{19}$$

$$\overline{ES}(X_1) = (\bar{S}(X_1), \bar{S}(\bar{X}_2), \bar{S}(\bar{X}_3), \ldots, \bar{S}(\bar{X}_T)) \tag{20}$$

where X_1 represents the starting numbers of aircraft at stage 1 and the X_t and $\bar{X}_t$ are defined recursively by

$$\begin{array}{lll} X_2 = Z_{ij}(X_1) & \qquad & \bar{X}_2 = Z_{k\ell}(X_1) \\ X_3 = Z_{ij}(X_2) & & \bar{X}_3 = Z_{k\ell}(\bar{X}_2) \\ \vdots & & \vdots \\ X_T = Z_{ij}(X_{T-1}) & & \bar{X}_T = Z_{k\ell}(\bar{X}_{T-1}) \end{array} \tag{21}$$

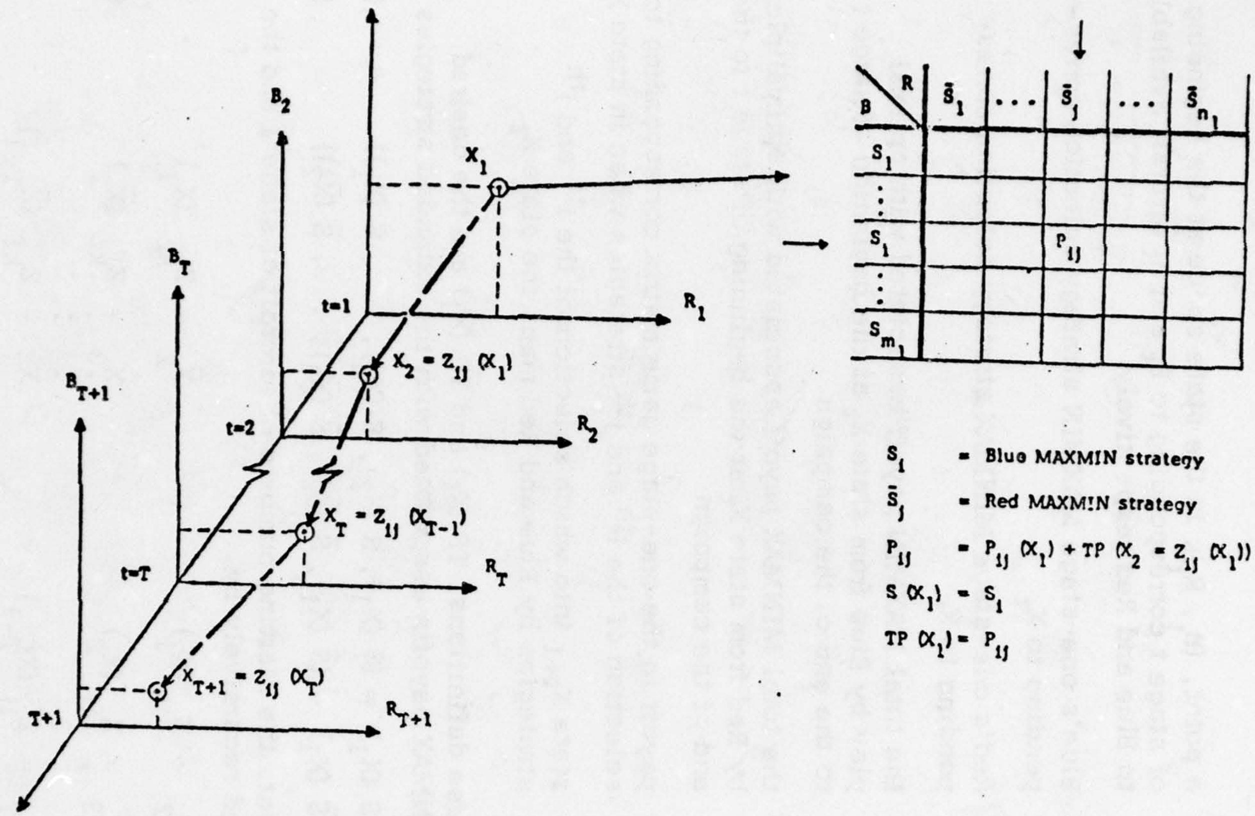


FIGURE 8

DYNAMIC PROGRAMMING SOLUTION FOR MAXMIN STRATEGIES AND BOUNDS

with i = index of Blue's MAXMIN strategy for the current stage
j = index of Red's MAXMIN strategy for the current stage
k = index of Blue's MINMAX strategy for the current stage
ℓ = index of Red's MINMAX strategy for the current stage

The dynamic programming approach for solving for TP (X_1), ES (X_1), $\overline{TP}$ (X_1), and $\overline{ES}$ (X_1) separates the solution for the MAXMIN total payoff and extended strategy from the MINMAX counterparts. Although all descriptions which follow concentrate on the MAXMIN solution, the MINMAX solution technique is exactly analogous.

<u>Idealized Approach</u>

To compute TP (X_1) and ES (X_1) using a general dynamic programming procedure, one begins at the beginning of the last stage of the campaign and computes one-stage MAXMIN payoffs and strategies, TP (X_T) and S (X_T), for all possible states X_T. The i, j[th] payoff, P_{ij}, in the MAXMIN game matrix for each state X_T is given by

$$P_{ij} = P_{ij} (X_T) + TP (X_{T+1} = Z_{ij} (X_T)) \qquad (22)$$

where P_{ij} (X_T) is the one-stage payoff and TP $(X_{T+1} = Z_{ij} (X_T))$ is the contribution of undamaged aircraft at the end of the war, B_{T+1} and R_{T+1}, to the value of the overall objective function (Equation 2).

After TP (X_T) and S (X_T) have been computed for all states X_T, these values are stored and the process moves backward to the beginning of stage T-1. Using the TP (X_T) just computed, the elements in the payoff matrix for each stage X_{T+1} are now given by

$$P_{ij} = P_{ij} (X_{T-1}) + TP (X_T = Z_{ij} (X_{T-1})) \qquad (23)$$

where P_{ij} (X_{T-1}) is the payoff contribution of the current stage and TP $(X_T = Z_{ij} (X_{T-1}))$ is the MAXMIN payoff associated with optimal play thereafter. Once again, a Blue MAXMIN strategy is computed for each game matrix along with the cumulative MAXMIN payoff TP (X_{T-1}) and each is stored for all possible states X_{T-1}. Processing moves backward to the preceding stage in time, game matrices are generated for all states, etc.

Eventually, through this iterative process, strategies and payoffs can be generated for all stages and states from stage T back to the beginning of stage 1. Since X_1, the hypothesized starting resource level, is

a point in the state space at the beginning of stage 1, the solution is complete. TP (X_1) is available immediately, while the one-stage strategies which compose ES (X_1) can be retrieved from storage by tracing through the states using the attrition map Z_{ij}. Perhaps even more important, as a by-product of the dynamic programming approach, solutions for all other X_t are also available since they have been stored during the processing required to solve for X_1. By simply retrieving the stored results, solutions for all shorter campaigns less than T stages which begin with any number of aircraft on either side can be easily generated.

Problems

Precise implementation of the general dynamic programming algorithm as described above requires evaluation of one-stage game for every possible state of every possible stage. The number of such states in a reasonable sized game is huge, even for the case in which each side has only one aircraft type. For example, each side might reasonably start with 1000 aircraft, making the number of possible (B_t, R_t) pairs more than one million.

Intuitively, one would suspect states which differ by only a few aircraft would produce approximately the same solution. ATACM exploits this intuitive feel by imposing a descrete grid upon the state space at each stage in a manner analogous to that used in DYGAM (Reference 4). If a grid such as that shown in Figure 9 were imposed, a tractable approach would be to explicitly compute one-stage strategies and payoffs for only the discrete grid points using the dynamic programming procedure described above. Unfortunately, as shown in Figure 9, the Z_{ij} (X_t) for a grid point X_t would generally not lie on a grid point in the subsequent stage's state space. Since TP $(Z_{ij} (X_t))$ is necessary to compute the elements in the one-stage game matrix (Equation (22)), an approximation technique is required for computing payoffs associated with points not on the grid.

Possible Approximations

One possible approximation technique for computing TP $(X_{t+1} = Z_{ij} (X_t))$ would be to linearly interpolate using the explicit payoffs for grid points adjacent to X_{t+1}. It can be shown that as the grid becomes finer the strategies which would result from using linear interpolation would approach those produced by the idealized dynamic programming solution. Similarly, the payoffs produced would approach the idealized lower bound TP (X_1), but not necessarily from below. In other words, although the

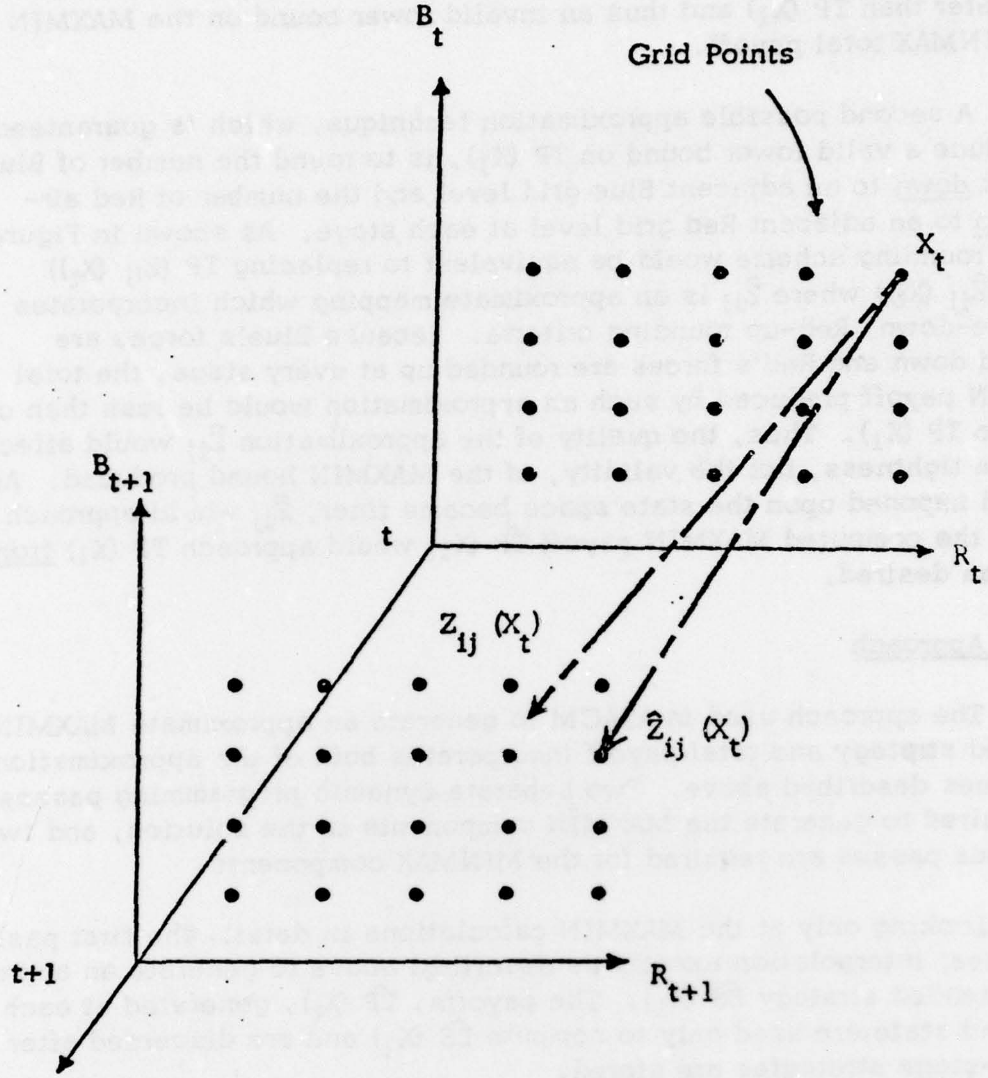


FIGURE 9

DYNAMIC PROGRAMMING APPROXIMATIONS

approximate extended strategy $\hat{ES}$ (X_1) produced by interpolation would intuitively be a good estimate of ES (X_1), the associated $\hat{TP}$ (X_1) might be greater than TP (X_1) and thus an invalid lower bound on the MAXMIN vs. MINMAX total payoff.

A second possible approximation technique, which is guaranteed to produce a valid lower bound on TP (X_1), is to round the number of Blue aircraft down to an adjacent Blue grid level and the number of Red aircraft up to an adjacent Red grid level at each stage. As shown in Figure 9, such a rounding scheme would be equivalent to replacing TP $(Z_{ij}\ (X_t))$ by TP $(\hat{Z}_{ij}\ (X_t))$ where $\hat{Z}_{ij}$ is an approximate mapping which incorporates the Blue-down, Red-up rounding criteria. Because Blue's forces are rounded down and Red's forces are rounded up at every stage, the total MAXMIN payoff produced by such an approximation would be less than or equal to TP (X_1). Thus, the quality of the approximation $\hat{Z}_{ij}$ would affect only the tightness, not the validity, of the MAXMIN bound produced. As the grid imposed upon the state space became finer, $\hat{Z}_{ij}$ would approach Z_{ij} and the computed MAXMIN payoff $\hat{TP}$ (X_1) would approach TP (X_1) from below as desired.

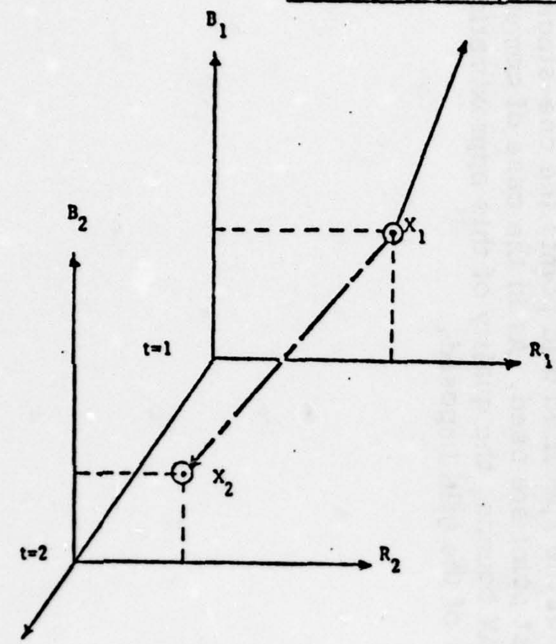
ATACM Approach

The approach used in ATACM to generate an approximate MAXMIN extended strategy and total payoff incorporates both of the approximation techniques described above. Two separate dynamic programming passes are required to generate the MAXMIN components of the solution, and two analogous passes are required for the MINMAX components.

Looking only at the MAXMIN calculations in detail, the first pass uses linear interpolation exactly as described above to generate an approximate extended strategy $\hat{ES}$ (X_1). The payoffs, $\hat{TP}$ (X_t), generated at each stage and state are used only to compute $\hat{ES}$ (X_1) and are discarded after the one-stage strategies are stored.

The second pass, designed to produce a lower bound on the MAXMIN total payoff TP (X_1), uses the $\hat{S}$ (X_t) stored during the first pass as fixed Blue strategies against which Red performs a MINMAX pass using the rounding scheme described above. As shown in Figure 10, in this second pass the game matrix for each stage and state has only one Blue strategy. Column maximums in these one-stage games are the single one-stage payoffs corresponding to each Blue strategy, and the MINMAX pass reduces to the calculation of the minimum payoff Red can achieve against $\hat{ES}$ (X_1).

B \ R	$\bar{S}_1$	$\bar{S}_2$	$\bar{S}_3$	...	$\bar{S}_{n_1}$
$\hat{S}(X_1)$	P_{11}	P_{12}	P_{13}	...	P_{1n_1}



STAGED GAME

OR

B \ R	$\overline{ES}_1$	$\overline{ES}_2$	$\overline{ES}_3$	...	$\overline{ES}_N$
$\hat{ES}(X_1)$	TP_{11}	TP_{12}	TP_{13}	...	TP_{1N}

EXTENDED GAME

FIGURE 10

ALTERNATIVE REPRESENTATIONS
OF THE SECOND DYNAMIC PROGRAMMING
PASS FOR COMPUTING $\hat{TP}(X_1)$

To see that this second MINMAX pass produces a lower bound on TP (X_1), consider two cases.

- If $\hat{ES}$ (X_1) is the same as ES (X_1), and no Blue-down/ Red-up rounding is required, the best Red can do in the MINMAX pass is to generate Red's MAXMIN extended strategy, which, by definition, produces a total payoff exactly equal to TP (X_1). If rounding is required, the total payoff would clearly be less than or equal to TP (X_1).

- If $\hat{ES}$ (X_1) is not equal to ES (X_1), and rounding is not required, the minimum total payoff Red can achieve in the MINMAX pass is the row minimum corresponding to $\hat{ES}$ (X_1) in the extended game of Figure 10. Since TP (X_1) is defined as the maximum over all row minimums, TP (X_1) must be bounded from below by the MINMAX payoff. As before, if rounding is required the value of the MINMAX payoff can only be reduced.

Finally, after two analogous dynamic programming passes for computing estimates of Red's MINMAX extended strategy $\overline{ES}$ (X_1) and the upper objective function bound TP (X_1), the solution is virtually complete. All that remains is the estimation of the MAXMIN vs. MINMAX payoff for starting state X_1 which is accomplished by actually playing the one-stage strategies stored for Blue and Red at each grid point. If during this calculation a Z_{ij} (X_t) does not fall on one of the specified grid points the one-stage strategies stored for the nearest point are used. As in the case of computing the MAXMIN and MINMAX bounds, the quality of this approximation depends upon the coarseness of the grid imposed.

REFERENCES

1. R. J. Galiano and F. A. Miercort, "Results of a Survey of Tactical Air Campaign Models," Contract No. ACDA/PAB-249, KETRON, Inc., November 1974.

2. Jerome Bracken, "Two Optimal Sortie Allocation Models," Paper P-992, Institute for Defense Analyses, December 1973.

3. S. Bonder, "VECTOR-O, The Battle Model Prototype," WSEG Report 222, Weapons System Evaluation Group, December 1973.

4. Control Analysis Corporation, "Development of an Algorithm to Solve Multi-Stage Games," 24 May 1973.